STIC-Biotech/ChemLib 160342

From: Chan, Christina
Sent: Monday, July 25, 2005 2:14 PM
To: Davis, Minh-Tam; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 10/019071

CRFE

Please rush. Thanks Chris

-----Original Message-----
From: **Davis, Minh-Tam**
Sent: Monday, July 25, 2005 2:11 PM
To: Chan, Christina
Subject: Rush search request for 10/019071

Please search in commercial database, issued patent files and PGPUB:
Amino acids 263-793 of SEQ ID NO:3.
Thank you.
MINH TAM DAVIS
ART UNIT 1642, ROOM 3A24, MB 3C18
272-0830

****
STAFF USE ONLY

Searcher:_____
Searcher Phone: 2-
Date Searcher Picked up:____
Date Completed:_____
Searcher Prep/Rev. Time:___
Online Time:_____

****
Type of Search

NA#:_____ AA#:_____
Interference:____ SPDI:___
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure#:_____ Text:___
Inventor:____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other(Specify):_____

O'Bryen, Barbara

From: Davis, Minh-Tam
Sent: Monday, July 25, 2005 3:50 PM
To: O'Bryen, Barbara
Subject: RE: problem with search request for 10/019071

Hi Barb,
amino acids 263-793 of Seq Id No **2, not 3.**
I apologize for the mistake.
Thanks
Tam

-----Original Message-----
From: O'Bryen, Barbara
Sent: Monday, July 25, 2005 3:38 PM
To: Davis, Minh-Tam
Subject: problem with search request for 10/019071
Importance: High

Hi Tam,
You had requested that we search "amino acids 263-793 of Seq Id No 3" from case 10/019071.
Seq ID 3 of this case is a nucleotide sequence, 45 nucleotides long. How would you like to modify your search request?
Please send your response directly to me, and **not** to the STIC-biotech/chemLib mailbox, since I now have the search on my desk.
Thanks,
Barb

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:54:32 ; Search time 42 Seconds
(without alignments)
1216.455 Million cell updates/sec

Title: US-10-019-071-2_COPY_263_793
Perfect score: 2896
Sequence: 1 DDSLNDCRIIFVDEVFKIER.....VNQPLQTVLNQLFPGYGNGR 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	600	20.7	641	2	E96612	probable transcrip
2	556	19.2	622	2	H96684	probable RING zinc
3	533.5	18.4	598	2	A96685	probable RING zinc
4	473	16.3	650	2	D96612	hypothetical prote
5	427	14.7	461	2	T01825	hypothetical prote
6	353.5	12.2	432	2	T00949	hypothetical prote
7	311.5	10.8	299	2	C75384	conserved hypothet
8	270.5	9.3	794	2	D84765	similar to mammali
9	263	9.1	788	2	C84616	similar to mammali
10	229.5	7.9	650	2	T06648	hypothetical prote
11	226	7.8	651	2	F84743	similar to mammali
12	194.5	6.7	669	2	F96756	hypothetical prote
13	192.5	6.6	954	2	G86312	hypothetical prote
14	170	5.9	4957	2	T03455	ALR protein - huma
15	170	5.9	5262	2	T03454	ALR protein - huma
16	162	5.6	1829	2	T34239	hypothetical prote
17	158	5.5	1280	2	T51500	hypothetical prote
18	157	5.4	811	2	T08738	hypothetical prote
19	153	5.3	1250	2	T00454	hypothetical prote
20	152.5	5.3	202	2	F88469	protein C28H8.9 [i
21	151.5	5.2	1350	2	T42697	hypothetical prote
22	150.5	5.2	728	2	S57142	hypothetical prote
23	150.5	5.2	1787	2	T20160	hypothetical prote
24	148	5.1	312	2	G84472	hypothetical prote
25	147	5.1	1722	1	I78879	retinoblastoma bin
26	146	5.0	429	2	C84640	similar to mammali
27	146	5.0	609	2	A43906	nuclear phosphopro
28	146	5.0	1479	2	T17401	transcription regu
29	143.5	5.0	1518	2	D96660	protein F2K11.14 [
30	139	4.8	449	2	T12495	hypothetical prote
31	138.5	4.8	530	2	I38558	Mi-2 autoantigen 2
32	138	4.8	1560	2	I54361	SMCX protein - hum
33	137.5	4.7	371	2	A55302	probable transcrip
34	137	4.7	1146	2	H96796	hypothetical prote
35	136	4.7	1033	2	I48775	Smcx protein (esca
36	134.5	4.6	796	2	T08555	pathogenesis-relat
37	133	4.6	571	2	T40911	probable PHD-type
38	130.5	4.5	397	2	S26731	neuro-D4 protein -
39	129.5	4.5	564	2	I48776	spermatogenesis re
40	129.5	4.5	2561	2	T24864	hypothetical prote
41	129	4.5	1088	2	T14917	homeotic protein P
42	127	4.4	1576	2	S65774	homeotic protein H
43	125.5	4.3	350	2	G01950	hypothetical prote
44	125	4.3	1257	2	T01020	hypothetical prote
45	124.5	4.3	2447	2	T16870	hypothetical prote

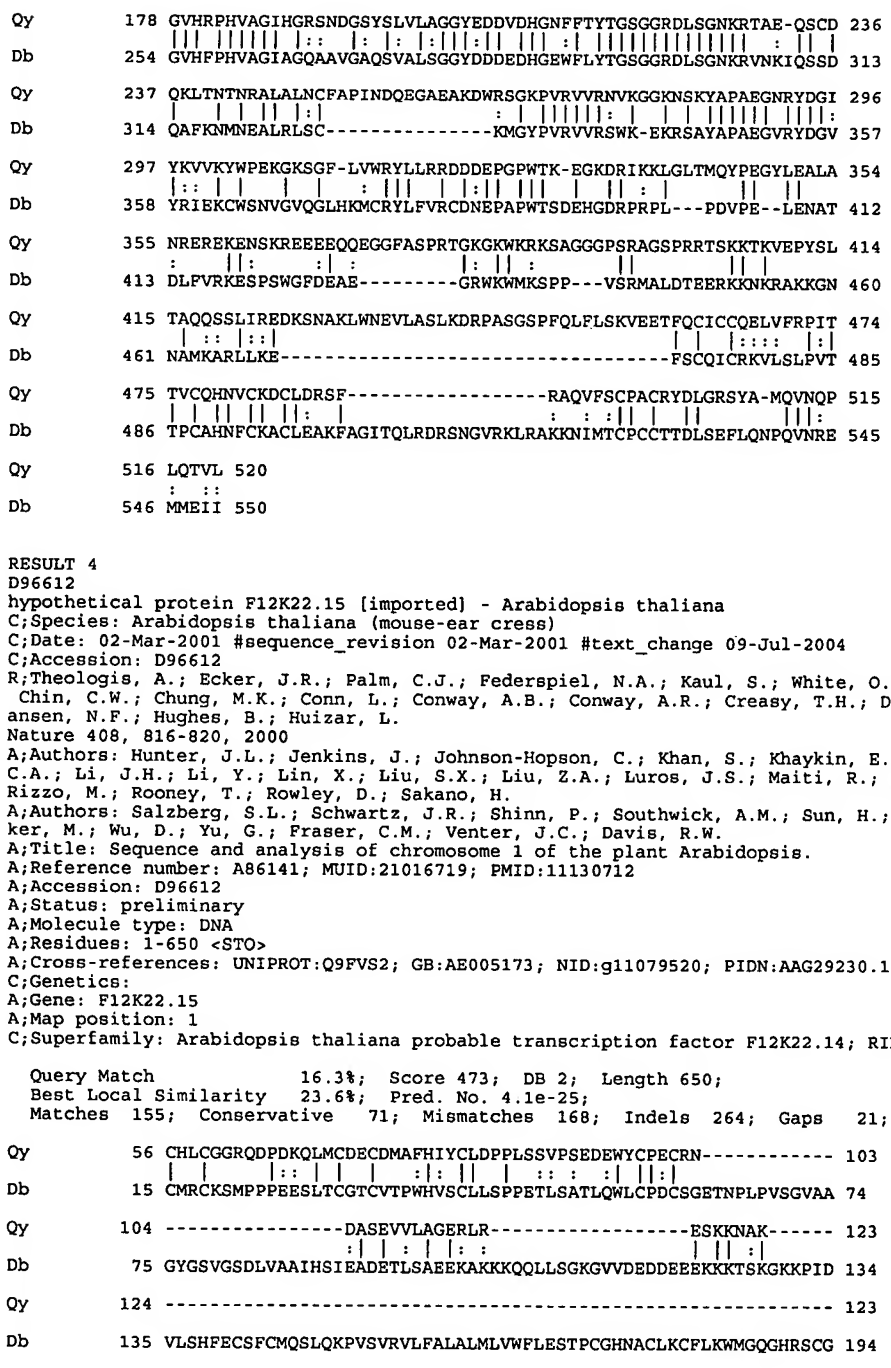
ALIGNMENTS

RESULT 1
E96612
probable transcription factor F12K22.14 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96612
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96612
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-641 <STO>
A;Cross-references: UNIPROT:Q9FVS3; GB:AE005173; NID:g11079528; PIDN:AAG29238.1; GSPDB:G
C;Genetics:
A;Gene: F12K22.14
A;Map position: 1
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 20.7%; Score 600; DB 2; Length 641;
Best Local Similarity 26.5%; Pred. No. 6.4e-34;
Matches 172; Conservative 68; Mismatches 160; Indels 248; Gaps 19;

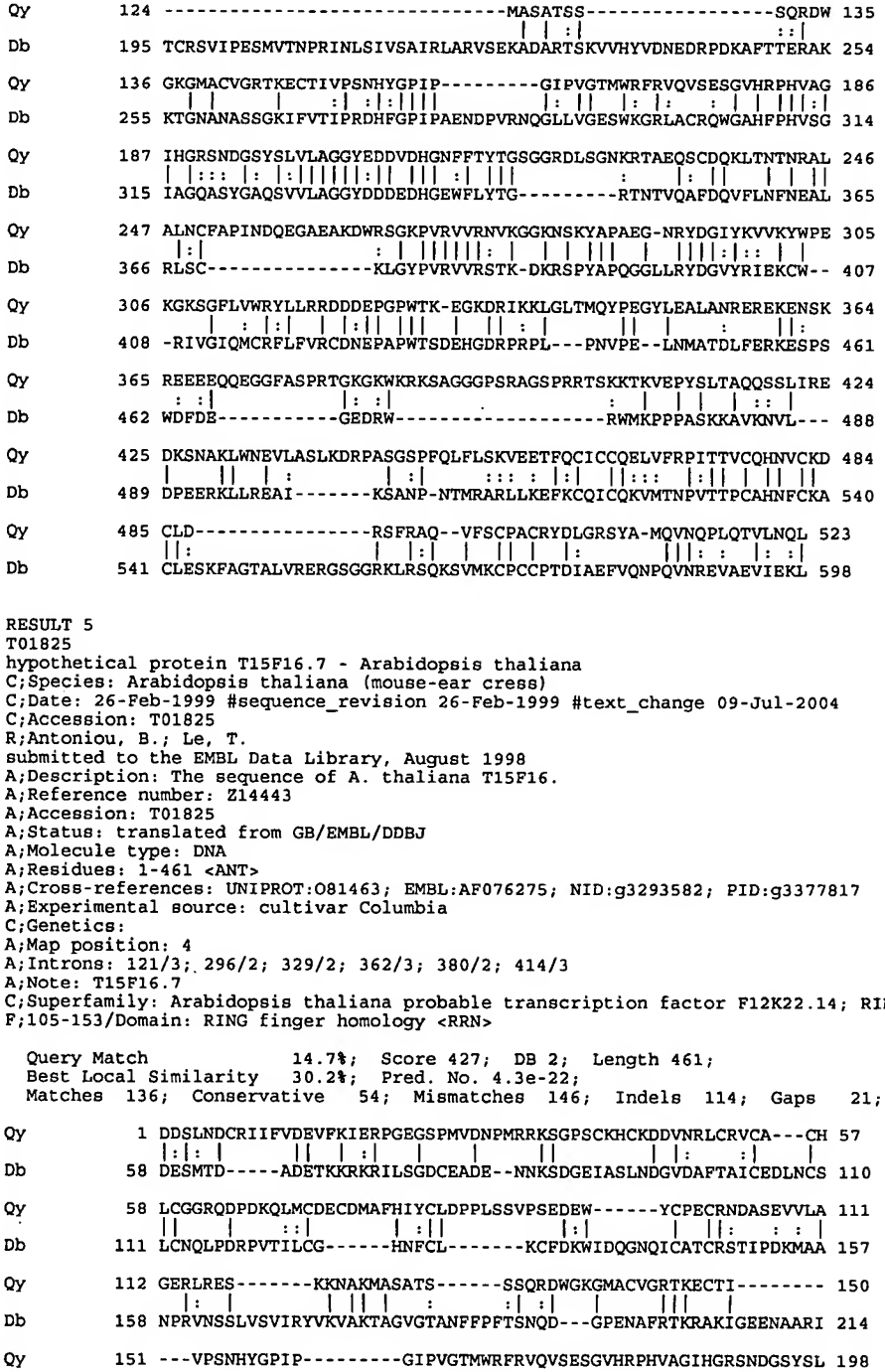
```
Qy          56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG--- 112
               |  |    |:: | |  |  :|: ||  |  :: |  :|:||:|  :   : ::|
Db          15 CMRCKSNPPPEESLTCGTCVTPWHVSCLSSPPKTLASTLQWHCPDCSGEIDPLPVSGGAT 74

Qy         113 ----------------------------------------------ERLRESK-------- 119
                                                            |  |: |
Db          75 GFESAGSDLVAAIRAIEADESLSTEEKAKMRQRLLSGKGVEEDDEEEKRKKKGKGKNPNL 134

Qy         120 ------------------------------------------------------------ 119

Db         135 DVLSALGDNLMCSFCMQLPERPVTKPCGHNACLKCFEKWMGQGKRTCGKCRSIIPEKMAK 194

Qy         120 ---------------KNAKMASATSS------SQRD------------WGKGMACVGRT 145
                              | :| |:||:|      | :|             ||  |  |
Db         195 NPRINSSLVAAIRLAKVSKSAAATTSKVFHFISNQDRPDKAFTTERAKKTGKANAASG-- 252

Qy         146 KECTIVPSNHYGPIP---------GIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSY 196
               |    :| :|:||||         |: ||  |  |::  : | | |||||| |:|  |:
Db         253 KIYVTIPPDHFGPIPAENDPVRNQGLLVGESWEDRLECRQWGAHFPHVAGIAGQSTYGAQ 312

Qy         197 SLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRT-AEQSCDQKLTNTNRALALNCFAPIN 255
```

```
Db        313 SVALSGGYKDDEDHGEWFLYTGSGGRDLSGNKRTNKEQSFDQKFEKSNAALKLSC----- 367

Qy        256 DQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWR 315

Db        368 ----------KLGYPVRVVRSHK-EKRSAYAPEEGVRYDGVYRIEKCW---RKVGVQVCR 413

Qy        316 YLLRRDDDEPGPWTK-EGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGG 374

Db        414 YLFVRCDNEPAPWTSDENGDRPRPI---PNIPE--LNMATDLFERKETPSWDFDE----- 463

Qy        375 FASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNE 434

Db        464 ------GEGCWKWM----------KPPPASKKS----------VNVLAPEERKNLR---- 493

Qy        435 VLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSF---- 490

Db        494 -------KAIKAAHSNTMRARLLKEFKCQICQQVLTLPVTTPCAHNFCKACLEAKFAGKT 546

Qy        491 --------------RAQVFSCPACRYDLGRSYA-MQVNQPLQTVLNQL 523

Db        547 LVRERSTGGRTLRSRKNVLNCPCCPTDISDFLQNPQVNREVAEVIEKL 594


RESULT 2
H96684
probable RING zinc finger protein F15E12.8 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H96684
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-622 <STO>
A;Cross-references: UNIPROT:Q9C8E1; GB:AE005173; NID:g11038468; PIDN:AAG27747.1; GSPDB:G
C;Genetics:
A;Gene: F15E12.8
A;Map position: 1
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

  Query Match             19.2%;  Score 556;  DB 2;  Length 622;
  Best Local Similarity   26.2%;  Pred. No. 7e-31;
  Matches  165;  Conservative   60;  Mismatches  171;  Indels  234;  Gaps   20;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECR------------ 102

Db         15 CMRCQVTPPSEETLTCGTCVTPWHVSCLLPE-SLASSTGDWECPDCSGVVVPSAAPGTGI 73

Qy        103 ---NDASEVVLAGER-------LRESKKNAKMASATSSSQRD----------------- 134

Db         74 SGPESSGSVLVAAIRAIQADVTLTEAEKAKKRQRLMSGGGDDGVDDEEKKKLEIFCSICI 133

Qy        135 ----------------------WGKG------MAC----------------------VG 143

Db        134 QLPERPVTTPCGHNFCLKCFEKWAVGQGKLTCMICRSKIPRHVAKNPRINLALVSAIRLA 193

Qy        144 RTKECT-------------------------------------------IVPSNHYGPIP- 160

Db        194 NVTKCSGEATAAKVHHIIRNQDRPDKAFTTERAVKTGKANAASGKFFVTIPRDHFGPIPA 253

Qy        161 --------GIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGN 212

Db        254 ANDVTRNQGVLVGESWEDRQECRQWGVHFPHVAGIAGQAAVGAQSVALSGGYDDDEDHGE 313

Qy        213 FFTYTGSGGRDLSGNKRTAE-QSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPV 271

Db        314 WFLYTGSGGRDLSGNKRVNKIQSSDQAFKNMNEALRLSC----------------KMGYPV 358

Qy        272 RVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGF-LVWRYLLRRDDDEPGPWTK 330

Db        359 RVVRSWK-EKRSAYAPAEGVRYDGVYRIEKCWSNVGVQGLHKMCRYLFVRCDNEPAPWTS 417

Qy        331 -EGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKS 389

Db        418 DEHGDRPRPL---PDVPE--LENATDLFVRKESPSWGFDEAE---------GRWKWMKSP 463

Qy        390 AGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPF 449

Db        464 P---VSRMALDTEERKKNKRAKKGNNAMKARLLKE------------------------ 495

Qy        450 QLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSF-----------------R 491

Db        496 ----------FSCQICRKVLSLPVTTPCAHNFCKACLEAKFAGITQLRDRSNGVRKLRAK 545

Qy        492 AQVFSCPACRYDLGRSYA-MQVNQPLQTVL 520

Db        546 KNIMTCPCCTTDLSEFLQNPQVNREMMEII 575


RESULT 3
A96685
probable RING zinc finger protein F15E12.5 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A96685
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96685
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-598 <STO>
A;Cross-references: UNIPROT:Q9C8E0; GB:AE005173; NID:g11038479; PIDN:AAG27758.1; GSPDB:G
C;Genetics:
A;Gene: F15E12.5
A;Map position: 1
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

  Query Match             18.4%;  Score 533.5;  DB 2;  Length 598;
  Best Local Similarity   26.8%;  Pred. No. 2.4e-29;
  Matches  162;  Conservative   54;  Mismatches  180;  Indels  209;  Gaps   19;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRN----------- 103

Db         15 CMRCQVNPPSEETLTCGTCVTPWHVSCLLPE-SLASSTGDWECPDCSGVVVPSAAPGTGI 73

Qy        104 -----------DASEVVLAGERLRESKKNAKMASATSSSQRD----------------- 134

Db         74 SGPESSGSVLVTAIRAIQADVTLTEAEKAKKRQRLMSGGGDDGVDDEEKKKLEIFCSICI 133

Qy        135 ----------------------WGKG------MACVGR---------------------- 144

Db        134 QLPERPVTTPCGHNFCLKCFEKWAVGQGKLTCMICRSKIPRHVAKNPRINLALVSAIRLA 193

Qy        145 --TK---ECTIVPSNHY----------------------GPIPGIPVGTMWRFRVQVSES 177

Db        194 NVTKCSGEATAAKVHHIIRNQDRPDKAFTTERAVKTGKANAASGVLVGESWEDRQECRQW 253
```



RESULT 4
D96612
hypothetical protein F12K22.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96612
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96612
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-650 <STO>
A;Cross-references: UNIPROT:Q9FVS2; GB:AE005173; NID:g11079520; PIDN:AAG29230.1; GSPDB:G
C;Genetics:
A;Gene: F12K22.15
A;Map position: 1
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger

Query Match 16.3%; Score 473; DB 2; Length 650;
Best Local Similarity 23.6%; Pred. No. 4.1e-25;
Matches 155; Conservative 71; Mismatches 168; Indels 264; Gaps 21;



RESULT 5
T01825
hypothetical protein T15F16.7 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01825
R;Antoniou, B.; Le, T.
submitted to the EMBL Data Library, August 1998
A;Description: The sequence of A. thaliana T15F16.
A;Reference number: Z14443
A;Accession: T01825
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <ANT>
A;Cross-references: UNIPROT:O81463; EMBL:AF076275; NID:g3293582; PID:g3377817
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 121/3; 296/2; 329/2; 362/3; 380/2; 414/3
A;Note: T15F16.7
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger
F;105-153/Domain: RING finger homology <RRN>

Query Match 14.7%; Score 427; DB 2; Length 461;
Best Local Similarity 30.2%; Pred. No. 4.3e-22;
Matches 136; Conservative 54; Mismatches 146; Indels 114; Gaps 21;

```
                || :|:|||| |: || | ||: : ||| |||: | |: : |: |:
Db        215 YVTVPFDHFGPIPAEHDPVRNQGVLVGESWENRVECRQWGVHLPHVSCIAGQEDYGAQSV 274

Qy        199 VLAGGYEDDVDHGNFFTYTG-SGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQ 257
              |::|||:|| ||| :| ||| | ||  :           ||:  : | || ::|
Db        275 VISGGYKDDEDHGEWFLYTGRSRGRHFANE--------DQEFEDLNEALRVSC------- 319

Qy        258 EGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYL 317
                    | |||||: |  : | ||| || ||||:|:: | |    |: | | |||
Db        320 --------EMGYPVRVVRSYK-DRYSAYAPKEGVRYDGVYRIEKCW---RKARFPVCRYL 367

Qy        318 LRRDDDEPGPW-TKEGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFA 376
                | |:|| || : |  || : |      ||  ||  ::    ||:   : :| :
Db        368 FVRCDNEPAPWNSDESGDRPRPL---PNIPE--LETASDLFERKESPSWDFDEAE----- 417

Qy        377 SPRTGKGKWKRKSAGGGPSRAGSPRRTSKK 406
                  |: :| :      | |   :|   |
Db        418 ----GRWRWMK------PPPANHEQRERMK 437


RESULT 6
T00949
hypothetical protein T3F12.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00949
R;Gnoj, L.; Huang, E.N.; Habermann, K.; Hameed, A.; Hasegawa, A.; Jensen, K.; Schutz, K.
R.; McCombie, W.R.
submitted to the EMBL Data Library, October 1997
A;Description: Arabidopsis thaliana BAC T3F12 from chromosome IV.
A;Reference number: Z14210
A;Accession: T00949
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-432 <GNO>
A;Cross-references: UNIPROT:O22280; EMBL:AC002983; NID:g2443899; PID:g2565009
C;Genetics:
A;Map position: 4
A;Introns: 121/3; 296/2; 329/3; 351/2; 385/3
A;Note: T3F12.10
C;Superfamily: Arabidopsis thaliana probable transcription factor F12K22.14; RING finger
F;105-153/Domain: RING finger homology <RRN>

  Query Match             12.2%;  Score 353.5;  DB 2;  Length 432;
  Best Local Similarity   26.9%;  Pred. No. 5e-17;
  Matches  121;  Conservative   49;  Mismatches  137;  Indels  143;  Gaps   19;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCA---CH 57
              |:|: |      ||  | :|   |     |      ||         | |:    :|    |
Db         58 DESMTD-----ADETKKRKRILSGDCEADE--NNKSDGEIASLNDGVDAFTAICEDLNCS 110

Qy         58 LCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEW------YCPECRNDASEVVLA 111
              ||     |    ::|      | :||        |:|       |  ||:   : : |
Db        111 LCNQLPDRPVTILCG------HNFCL-------KCFDKWIDQGNQICATCRSTIPDKMAA 157

Qy        112 GERLRES-------KKNAKMASATS------SSQRDWGKGMACVGRTKECTI-------- 150
                |:  |        | || |   :      :| :|   |     |||   |
Db        158 NPRVNSSLVSVIRYVKVAKTAGVGTANFFPPTSNQD---GPENAFRTKRAKIGEENAARI 214

Qy        151 ---VPSNHYGPIP---------GIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSL 198
                 || :|:|||          |: ||  |  ||:  : ||| |||: | |: : |: |:
Db        215 YVTVPFDHFGPIPAEHDPVRNQGVLVGESWENRVECRQWGVHLPHVSCIAGQEDYGAQSV 274

Qy        199 VLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQE 258
              |::|||:|| ||| :| |||
Db        275 VISGGYKDDEDHGEWFLYTGR----------------------------------------- 295

Qy        259 GAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPE-KGKSGFLVWRYL 317
                    || |           : | ||| || ||||:|:: | | : :    | | |||
Db        296 -------RSYK----------DRYSAYAPKEGVRYDGVYRIEKCWRKARFPDSFKVCRYL 338

Qy        318 LRRDDDEPGPW-TKEGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFA 376
                | |:|| || : |  || : |      ||  ||  ::    ||:   : :| :

Db        339 FVRCDNEPAPWNSDESGDRPRPL---PNIPE--LETASDLFERKESPSWDFDEAE----- 388

Qy        377 SPRTGKGKWKRKSAGGGPSRAGSPRRTSKK 406
                  |: :| :      | |   :|   |
Db        389 ----GRWRWMK------PPPANHEQRERMK 408


RESULT 7
C75384
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: C75384
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: C75384
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <WHI>
A;Cross-references: UNIPROT:Q9RU61; GB:AE001997; GB:AE000513; NID:g6459292; PIDN:AAF1109
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1533
A;Map position: 1

  Query Match             10.8%;  Score 311.5;  DB 2;  Length 299;
  Best Local Similarity   38.2%;  Pred. No. 2.7e-14;
  Matches   65;  Conservative   34;  Mismatches   50;  Indels   21;  Gaps    3;

Qy        155 HYGPIPGIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFF 214
              |:| :||:  |  :  | :: ::||| |  ||| | :::|: |:||:|||||| | |:
Db          4 HFGAVPGVVPGMAFVNRQELRDAGVHLPTQAGISGSASEGADSIVLSGGYEDDRDEGDVI 63

Qy        215 TYTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVV 274
               ||| ||||      |  |    |:|   | |||::                | | |:||
Db         64 LYTGEGGRD----PLTGHQVKPQQLVRGNLALAIS--------------HRDGLPLRVT 104

Qy        275 RNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDE 324
              |  :   :|:::|  | :| |:|:|  :| | |:||||:||: | | :::
Db        105 RGHR--HSSQFSPQSGYQYAGLYRVDDHWREVGRSGFLIWRFRLTRLENQ 152


RESULT 8
D84765
similar to mammalian MHC III region protein G9a [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84765
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84765
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-794 <STO>
A;Cross-references: UNIPROT:O82175; GB:AE002093; NID:g3668088; PIDN:AAC61820.1; GSPDB:GN
C;Genetics:
A;Gene: At2g35160
A;Map position: 2

  Query Match             9.3%;  Score 270.5;  DB 2;  Length 794;
  Best Local Similarity   29.5%;  Pred. No. 5.6e-11;
  Matches   71;  Conservative   43;  Mismatches   92;  Indels   35;  Gaps    7;

Qy        103 NDASEVVLAGERLRESKK---NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPI 159
              ::  : ::  |  |  |:   | |:    |   :  || :             :   | :
```

```
Db        318 HETCKKIMQEEEARPRKRDGGNFKVVCEASKILKSKGKNL----------YSGTQIIGTV 367

Qy        160 PGIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDG----SYSLVLAGGYEDDVDHGNFFT 215
              ||: ||  :::|::::  |:||| :||    :||    : |:| :||| | :|: :
Db        368 PGVEVGDEFQYRMELNLLGIHRPSQSGIDYMKDDGGELVATSIVSSGGYNDVLDNSDVLI 427

Qy        216 YTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVR 275
              ||| ||    | |:  |   ||:|  | ||  :    || :                ||||:|
Db        428 YTGQGGN--VGKKKNNEPPKDQQLVTGNLALKNS----INKK----------NPVRVIR 470

Qy        276 NVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPG-PWTKEGKD 334
             :|       |:   |||:| | :|| | |  | ||::: |||   :|  || :  |
Db        471 GIKNTTLQSSVVAKNYVYDGLYLVEEYWEETGSHGKLVFKFKLRRIPGQPELPWKEVAKS 530

Qy        335 R 335
             :
Db        531 K 531


RESULT 9
C84616
similar to mammalian MHC III region protein G9a [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84616
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84616
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-788 <STO>
A;Cross-references: GB:AE002093; NID:g4314371; PIDN:AAD15582.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g22740
A;Map position: 2

  Query Match             9.1%;  Score 263;  DB 2;  Length 788;
  Best Local Similarity   30.2%;  Pred. No. 1.9e-10;
  Matches   80;  Conservative   49;  Mismatches   88;  Indels   48;  Gaps   13;

Qy        127 ATSSSQRDWGKGMACVGRTKECTIVPSN--------H-YGPIPGIPVGTMWRFRVQVSES 177
              |    ||  |||:      :  ||:  |        |  | :||: ||  :::|::::
Db        291 AKPEDQRRKGKGLRI--DFEASTILKRNGKFLNSGVHILGEVPGVEVGDEFQYRMELNIL 348

Qy        178 GVHRPHVAGI----HGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAE- 232
              |:|:|  |||    :|::   : |:| :|||:| :|: :  |||| ||  :   |:  |
Db        349 GIHKPSQAGIDYMKYGKAKVAT-SIVASGGYDDHLDNSDVLTYTGQGGNVMQVKKKGEEL 407

Qy        233 -QSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGN 291
              : .||||   | ||| :   |  |           ||||:|   |   | :  ::|
Db        408 KEPEDQKLITGNLALATS----IEKQ----------TPVRVIR---GKHKSTHDKSKGG 449

Qy        292 R--YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPG-PWTKEGKDRIKKLGLTMQYPEG 348
              |||:| | ||| : |  |  |::: |||   :|   | :  | : |        | ||
Db        450 NYVYDGLYLVEKYWQQVGSHGMNVFKFQLRRIPGQPELSWVEVKKSKSK-------YREG 502

Qy        349 Y--LEALANREREKENSKREEEEQQ 371
              |:    :|:   ::  | ::::
Db        503 LCKLDISEGKEQSPISAVNEIDDEK 527


RESULT 10
T06648
hypothetical protein T6G15.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06648
R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15791
A;Accession: T06648
A;Molecule type: DNA
A;Residues: 1-650 <BEV>
A;Cross-references: UNIPROT:Q9T0G7; EMBL:AL049656; GSPDB:GN00062; ATSP:T6G15.10
A;Experimental source: cultivar Columbia; BAC clone T6G15
C;Genetics:
A;Gene: ATSP:T6G15.10
A;Map position: 4

  Query Match             7.9%;  Score 229.5;  DB 2;  Length 650;
  Best Local Similarity   28.8%;  Pred. No. 3.1e-08;
  Matches   74;  Conservative   39;  Mismatches   83;  Indels   61;  Gaps   11;

Qy        110 LAGERLRESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWR 169
              : | |  |    |    |  ||    | : | |           | |||: || ::
Db        172 VGGRRARADGK----AGKAGSMMRDC---MLWMNRDKRIV-------GSIPGVQVGDIFF 217

Qy        170 FRVQVSESGVHRPHVAGIH----GRSNDG---SYSLVLAGGYEDDVDHGNFFTYTGSGGR 222
              || ::   |:|   :||       |::|   : |::::|||||| | |:   ||| ||:
Db        218 FRFELCVMGLHGHPQSGIDFLTGSLSSNGEPIATSVIVSGGYEDDDDQGDVIMYTGQGGQ 277

Qy        223 DLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKN 282
              |     |   |:  |:|   | |:  : :                 |  |||:| :
Db        278 D-----RLGRQAEHQRLEGGNLAMERSMY--------------YGIEVRVIRGL----- 312

Qy        283 SKYAPAEGNR---YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKL 339
               ||     :|  |||::::|  | : |||||  |::| | | : :       |    : |
Db        313 -KYENEVSSRVYVYDGLFRIVDSWFDVGKSGFGVFKYRLERIEGQ-----AEMGSSVLKF 366

Qy        340 GLTMQ------YPEGYL 350
                |::       | ||:
Db        367 ARTLKTNPLSVRPRGYI 383


RESULT 11
F84743
similar to mammalian MHC III region protein G9a [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84743
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84743
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-651 <STO>
A;Cross-references: UNIPROT:O22781; GB:AE002093; NID:g2459412; PIDN:AAB80647.1; GSPDB:GN
C;Genetics:
A;Gene: At2g33290
A;Map position: 2

  Query Match             7.8%;  Score 226;  DB 2;  Length 651;
  Best Local Similarity   26.4%;  Pred. No. 5.5e-08;
  Matches   85;  Conservative   55;  Mismatches  106;  Indels   76;  Gaps   17;

Qy         83 LDPP-------LSSVPSEDEWYCPECRNDASEVVLAG----ERLRESKKNAKMA------ 125
              ::||       :|:|    ::  |    : : : :|    : ||:  |  :|
Db        108 IEPPPGFKDNRVSTVVVSPKFERP---RELARIAILGHEQRKELRQVMKRTRMTYESLRI 164

Qy        126 --SATSSSQRDWGKG--------MACVGRTKECTIVPSNH-YGPIPGIPVGTMWRFRVQV 174
                | |       |:|          | : | :   :     |  ||: |: || :: :|:::
Db        165 HLMAESMKNHVLGQGRRRRSDMAAAYIMRDRGLWLNYDKHIVGPVTGVEVGDIPFYRMEL 224

Qy        175 SESGVHRPHVAGIH----GRSNDG---SYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGN 227
              |:|    |||      ||  |   : |:|::|||||| | |:   ||| ||:|
Db        225 CVLGLHGQTQAGIDCLTAERSATGEPIATSIVVSGGYEDDEDTGDVLVYTGHGGQD---- 280

Qy        228 KRTAEQSCD-QKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYA 286
```

```
Db        281 --HQHKQCDNQRLVGGNLGM---------------ERSMHYGIEVRVIRGI------KYE 317

Qy        287 PAEGNR---YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEP--GPWTKEGKDRIKKLGL 341
Db        318 NSISSKVYVYDGLYKIVDWWFAVGKSGFGVFKFRLVRIEGQPMMGSAVMRFAQTLRNKP- 376

Qy        342 TMQYPEGYLE-ALANREREKEN 362
Db        377 SMVRPTGYVSFDLSNK---KEN 395


RESULT 12
F96756
hypothetical protein F3N23.30 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: F96756
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96756
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-669 <STO>
A;Cross-references: GB:AE005173; NID:g5903099; PIDN:AAD55657.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3N23.30
A;Map position: 1

  Query Match             6.7%;  Score 194.5;  DB 2;  Length 669;
  Best Local Similarity   29.5%;  Pred. No. 8.6e-06;
  Matches   67;  Conservative   31;  Mismatches   92;  Indels   37;  Gaps    8;

Qy        157 GPIPGIPVGTMWRFRVQVSESGVHRPHVAGIH------GRSNDG-SYSLVLAGGYEDDVD 209
Db        208 GTVPGIEVGDIFFSRIEMCLVGLHMQTMAGIDYIISKAGSDEESLATSIVSSGRYEGEAQ 267

Qy        210 HGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGK 269
Db        268 DPESLIYSGQG-----GNADKNRQASDQKLERGNLAL---------------ENSLRKGN 307

Qy        270 PVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEP---G 326
Db        308 GVRVVR---GEEDAASKTGKIYIYDGLYSISESWVEKGKSGCNTFKYKLVRQPGQPPAFG 364

Qy        327 PWTKEGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEG 373
Db        365 FWKSVQK---WKEGLTTR-PGLILPDLTSGAESKPVSLVNDVDEDKG 407


RESULT 13
G86312
hypothetical protein F2H15.1 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C;Accession: G86312
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G86312
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-954 <STO>
A;Cross-references: GB:AE005172; NID:g9665056; PIDN:AAF97258.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match             6.6%;  Score 192.5;  DB 2;  Length 954;
  Best Local Similarity   30.6%;  Pred. No. 1.8e-05;
  Matches   55;  Conservative   28;  Mismatches   58;  Indels   39;  Gaps    6;

Qy        157 GPIPGIPVGTMWRFRVQVSESGVHRPHVAGI-------HGRSNDGSYSLVLAGGYEDDVD 209
Db        227 GAVPGIHVGDIFYYWGEMCLVGLHKSNYGGIDFFTAAESAVEGHAAMCVVTAGQYDGETE 286

Qy        210 HGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGK 269
Db        287 GLDTLIYSGQGGTDVYGNAR------DQEMKGGNLAL-----------EASVSK----GN 325

Qy        270 PVRVVRNVKGGKNSKYAPAEGNR----YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEP 325
Db        326 DVRVVRGV-------IHPHENNQKIYIYDGMYLVSKFWTVTGKSGFKEFRFKLVRKPNQP 378


RESULT 14
T03455
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03455
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03455
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-4957 <PRA>
A;Cross-references: UNIPROT:O14686; EMBL:AF010404; NID:g2358286; PIDN:AAC51735.1; PID:g2
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

  Query Match             5.9%;  Score 170;  DB 2;  Length 4957;
  Best Local Similarity   20.3%;  Pred. No. 0.0042;
  Matches  120;  Conservative   72;  Mismatches  221;  Indels  178;  Gaps   27;

Qy          1 DDSLNDCRIIF--VDEVFKIE----------RPGEGSPMVDNPMRRKSGPSCKHCKDDVN 48
Db        777 DDTMQNTVVLFSNTDKFVLMQDMCVVCGSFGRGAEGHLLACSQCSQCYHPYCVNSKITKV 836

Qy         49 RL-----CRVC-ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPS---------- 92
Db        837 MLLKGWRCVECIVCEVCGQASDPSRLLLCDDCDISYHTYCLDPPLLTVPKGGWKCKWCVS 896

Qy         93 ----------EDEWY-------------CPECRND-ASEVVLAGERLRESKKNAKMAS 126
Db        897 CMQCGAASPGFHCEWQNSYTHCGPCASLVTCPICHAPYVEEDLLIQCRHCERWMHAGCES 956

Qy        127 ATSSSQRDWG--KGMACVGRTKECTIVPSNHYGPIPGIPVGT------MWRFR-VQVSES 177
Db        957 LFTEDDVDHAPDEGFDCVS-CQPYVVKPVAPVAPPELVPMKVKEPEPQYFRFEGVWLTET 1015

Qy        178 G----------VHR----------PHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFT- 215
Db       1016 GMALLRNLTMSPLHKRRQRRGRLGLPGEAGLEGSEPSDALG-------PDDKKDGDLDTD 1068
```

```
Qy        216 --YTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPIN-----DQEGAEAKDWRSG 268
Db       1069 ELLKGEGG--------VEHMECEIKLE----------GPVSPDVEPGKEETEESKKRKR 1109

Qy        269 KPVR------VVRNVKGGKNSKYAP-AEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRD 321
Db       1110 KPYRPGIGGFMVRQRKSHTRTKKGPAAQAEVLSGDGQPDEVIPADLPAEGAVEQSL--AE 1167

Qy        322 DDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE----------------------- 357
Db       1168 GDEKKKQQRRGRKRSKLEGM---FPAYLQEAFFGKELLDLSRKALFAVGVGRPSFGLGTP 1224

Qy        358 ----------REKENSKR---------EEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAG 398
Db       1225 KAKGDGGSERKELPTSQKGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVPSDPEKPG 1284

Qy        399 SPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRP-ASGSP 448
Db       1285 TPGEGMLSSDLDRIS-TEELPKM--ESKDLQQLFKDVLGSEREQHLGCGTP 1332
```

RESULT 15
T03454
ALR protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03454
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
Oncogene 15, 549-560, 1997
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A;Reference number: Z14954; MUID:97388474; PMID:9247308
A;Accession: T03454
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5262 <PRA>
A;Cross-references: UNIPROT:O14686; EMBL:AF010403; NID:g2358284; PIDN:AAC51734.1; PID:g2
C;Genetics:
A;Gene: ALR
A;Map position: 12
C;Superfamily: acute lymphoblastic leukemia protein, ALR type
C;Keywords: alternative splicing

  Query Match   5.9%; Score 170; DB 2; Length 5262;
  Best Local Similarity   20.3%; Pred. No. 0.0044;
  Matches 120; Conservative   72; Mismatches 221; Indels 178; Gaps   27;

```
Qy          1 DDSLNDCRIIF--VDEVFKIE----------RPGEGSPMVDNPMRRKSGPSCKHCKDDVN 48
Db       1082 DDTMQNTVVLFSNTDKFVLMQDMCVVCGSFGRGAEGHLLACSQCSQCYHPYCVNSKITKV 1141

Qy         49 RL-----CRVC-ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPS---------- 92
Db       1142 MLLKGWRCVECIVCEVCGQASDPSRLLLCDDCDISYHTYCLDPPLLTVPKGGWKCKWCVS 1201

Qy         93 -----------EDEWY-------------CPECRND-ASEVVLAGERLRESKKNAKMAS 126
Db       1202 CMQCGAASPGFHCEWQNSYTHCGPCASLVTCPICHAPYVEEDLLIQCRHCERWMHAGCES 1261

Qy        127 ATSSSQRDWG--KGMACVGRTKECTIVPSNHYGPIPGIPVGT------MWRFR-VQVSES 177
Db       1262 LFTEDDVDHAPDEGFDCVS-CQPYVVKPVAPVAPPELVPMKVKEPEPQYFRFEGVWLTET 1320

Qy        178 G-----------VHR----------PHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFT- 215
Db       1321 GMALLRNLTMSPLHKRRQRRGRLGLPGEAGLEGSEPSDALG-------PDDKKDGDLDTD 1373

Qy        216 --YTGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPIN-----DQEGAEAKDWRSG 268
Db       1374 ELLKGEGG--------VEHMECEIKLE----------GPVSPDVEPGKEETEESKKRKR 1414

Qy        269 KPVR------VVRNVKGGKNSKYAP-AEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRD 321
Db       1415 KPYRPGIGGFMVRQRKSHTRTKKGPAAQAEVLSGDGQPDEVIPADLPAEGAVEQSL--AE 1472

Qy        322 DDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE----------------------- 357
Db       1473 GDEKKKQQRRGRKRSKLEGM---FPAYLQEAFFGKELLDLSRKALFAVGVGRPSFGLGTP 1529

Qy        358 ----------REKENSKR---------EEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAG 398
Db       1530 KAKGDGGSERKELPTSQKGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVPSDPEKPG 1589

Qy        399 SPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRP-ASGSP 448
Db       1590 TPGEGMLSSDLDRIS-TEELPKM--ESKDLQQLFKDVLGSEREQHLGCGTP 1637
```

Search completed: July 25, 2005, 17:07:33
Job time : 45 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 25, 2005, 17:00:57 ; Search time 158 Seconds
(without alignments)
1307.310 Million cell updates/sec

Title: US-10-019-071-2_COPY_263_793
Perfect score: 2896
Sequence: 1 DDSLNDCRIIFVDEVFKIER.....VNQPLQTVLNQLFPGYGNGR 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1741741 seqs, 388992284 residues

Total number of hits satisfying chosen parameters: 1741741

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2896	100.0	793	16	US-10-751-736-115	Sequence 115, App
2	2887	99.7	793	14	US-10-123-568-2	Sequence 2, Appli
3	2887	99.7	793	15	US-10-188-832-27	Sequence 27, Appl
4	2887	99.7	793	16	US-10-370-715B-456	Sequence 456, App
5	2815	97.2	780	15	US-10-295-027-156	Sequence 156, App
6	1718.5	59.3	802	16	US-10-476-924-7	Sequence 7, Appli
7	1297	44.8	645	15	US-10-126-103-113	Sequence 113, App
8	1297	44.8	645	15	US-10-431-096-113	Sequence 113, App
9	823.5	28.4	198	9	US-09-867-550-766	Sequence 766, App
10	736	25.4	136	14	US-10-123-568-3	Sequence 3, Appli
11	672	23.2	133	9	US-09-764-864-1301	Sequence 1301, Ap
12	626.5	21.6	694	15	US-10-425-114-39797	Sequence 39797, A
13	626.5	21.6	709	15	US-10-424-599-174378	Sequence 174378,
14	608	21.0	645	16	US-10-739-930-6375	Sequence 6375, Ap
15	588.5	20.3	617	15	US-10-225-066A-522	Sequence 522, App
16	588.5	20.3	617	15	US-10-374-780A-2314	Sequence 2314, Ap
17	498	17.2	750	16	US-10-437-963-156872	Sequence 156872,
18	478.5	16.5	774	16	US-10-437-963-180766	Sequence 180766,
19	420.5	14.5	178	9	US-09-764-864-1303	Sequence 1303, Ap
20	357.5	12.3	110	9	US-09-764-864-848	Sequence 848, App
21	331	11.4	438	15	US-10-424-599-263042	Sequence 263042,
22	279	9.6	684	16	US-10-437-963-132977	Sequence 132977,
23	271.5	9.4	543	16	US-10-425-115-333474	Sequence 333474,
24	266.5	9.2	273	15	US-10-425-114-41693	Sequence 41693, A
25	263	9.1	790	16	US-10-739-930-6821	Sequence 6821, Ap
26	245.5	8.5	767	16	US-10-425-115-322796	Sequence 322796,
27	245.5	8.5	810	15	US-10-425-114-54293	Sequence 54293, A
28	236	8.1	707	15	US-10-424-599-277747	Sequence 277747,
29	236	8.1	708	17	US-10-732-923-15064	Sequence 15064, A
30	235.5	8.1	670	17	US-10-732-923-15058	Sequence 15058, A
31	232.5	8.0	557	15	US-10-424-599-167978	Sequence 167978,
32	230.5	8.0	682	16	US-10-437-963-109882	Sequence 109882,
33	229.5	7.9	600	15	US-10-425-114-57455	Sequence 57455, A
34	229.5	7.9	856	15	US-10-310-154-600	Sequence 600, App
35	229.5	7.9	856	17	US-10-732-923-545	Sequence 545, App
36	224	7.7	297	15	US-10-424-599-171858	Sequence 171858,
37	223	7.7	678	16	US-10-437-963-145990	Sequence 145990,
38	221	7.6	736	17	US-10-732-923-15065	Sequence 15065, A
39	221	7.6	768	16	US-10-437-963-133503	Sequence 133503,
40	219.5	7.6	396	15	US-10-425-114-70940	Sequence 70940, A
41	219.5	7.6	426	16	US-10-425-115-267706	Sequence 267706,
42	219	7.6	250	16	US-10-767-701-33924	Sequence 33924, A
43	218	7.5	697	16	US-10-425-115-296869	Sequence 296869,
44	216.5	7.5	673	16	US-10-437-963-159376	Sequence 159376,
45	215.5	7.4	730	15	US-10-425-114-72673	Sequence 72673, A

ALIGNMENTS

RESULT 1
US-10-751-736-115
; Sequence 115, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10/751,736
; CURRENT FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIOR FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 115
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-751-736-115

Query Match 100.0%; Score 2896; DB 16; Length 793;
Best Local Similarity 100.0%; Pred. No. 1.7e-233;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 2
US-10-123-568-2
; Sequence 2, Application US/10123568
; Publication No. US20030194713A1
; GENERAL INFORMATION:
;  APPLICANT: Hitoshi, Yasumichi
;  APPLICANT:  Jenkins, Yonchu
;  APPLICANT:  Rigel Pharmaceuticals, Inc.
;  TITLE OF INVENTION: NP95:  Methods of Assaying for Cell Cycle Modulators
;  FILE REFERENCE: 021044-003400US
;  CURRENT APPLICATION NUMBER: US/10/123,568
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human NP95 nuclear zinc finger protein
US-10-123-568-2

  Query Match            99.7%;  Score 2887;  DB 14;  Length 793;
  Best Local Similarity  99.6%;  Pred. No. 9.8e-233;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 3
US-10-188-832-27
; Sequence 27, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
;  APPLICANT: Mack, David H.
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
;  TITLE OF INVENTION:  and Methods of Screening for Modulators of Bladder
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 018501-002330US
;  CURRENT APPLICATION NUMBER: US/10/188,832
;  CURRENT FILING DATE:  2002-11-22
;  PRIOR APPLICATION NUMBER: US 60/302,814
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: US 60/310,099
;  PRIOR FILING DATE: 2001-08-03
;  PRIOR APPLICATION NUMBER: US 60/343,705
;  PRIOR FILING DATE: 2001-11-08
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/372,246
;  PRIOR FILING DATE: 2002-04-12
;  NUMBER OF SEQ ID NOS: 207
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-188-832-27

  Query Match            99.7%;  Score 2887;  DB 15;  Length 793;
  Best Local Similarity  99.6%;  Pred. No. 9.8e-233;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502
```

```
Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 4
US-10-370-715B-456
; Sequence 456, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
;    Patin Docket Preview
;  APPLICANT: BODARY,SARAH C.
;  APPLICANT:  CLARK,HILLARY
;  APPLICANT:  BRISDELL,HUNTE
;  APPLICANT:  JACKMAN,JANET
;  APPLICANT:  SCHOENFELD,JILL R.
;  APPLICANT:  WILLIAMS,P. MICKEY
;  APPLICANT:  WOOD,WILLIAM I.
;  APPLICANT:  WU,THOMAS D.
;  TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
;  TITLE OF INVENTION:  Related Diseases
;  FILE REFERENCE: P1948R1-US
;  CURRENT APPLICATION NUMBER: US/10/370,715B
;  CURRENT FILING DATE:  2003-02-21
;  NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 456
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-370-715B-456

  Query Match             99.7%;  Score 2887;  DB 16;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 9.8e-233;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 5
US-10-295-027-156
; Sequence 156, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 156
;   LENGTH: 780
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-156

  Query Match             97.2%;  Score 2815;  DB 15;  Length 780;
  Best Local Similarity   99.6%;  Pred. No. 1e-226;
  Matches  516;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
                |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 518
                ||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 780
```

```
RESULT 6
US-10-476-924-7
; Sequence 7, Application US/10476924
; Publication No. US20040152093A1
; GENERAL INFORMATION:
;  APPLICANT: YUE, Henry; DING, Li;
;  APPLICANT:  BAUGHN, Mariah R.; LAL,Preeti G.;
;  APPLICANT:  YUE, Huibin; HAFALIA, April J.A.;
;  APPLICANT:  LEE, Ernestine A.; ISON, Craig H.;
;  APPLICANT:  BECHA, Shanya D.; GURURAJAN, Rajagopal;
;  APPLICANT:  EMERLING, Brooke M.; GRIFFIN, Jennifer A.;
;  APPLICANT:  TANG, Y. Tom; LU, Dyung Aina M.;
;  APPLICANT:  YAO, Monique G.; CHAWLA, Narinder K.;
;  APPLICANT:  RAMKUMAR, Jayalaxmi; GANDHI, Ameena R.;
;  APPLICANT:  LEE, Soo Yeun; RICHARDSON, Thomas W.;
;  APPLICANT:  YANG, Junming; ELLIOTT, Vicki S.;
;  APPLICANT:  LU, Yan; THANGAVELU, Kavitha;
;  APPLICANT:  HE, Ann; AZIMZAI, Yalda;
;  APPLICANT:  RAUMANN, Brigette E.; SWARNAKAR, Anita;
;  APPLICANT:  BURFORD, Neil
;  TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
;  FILE REFERENCE: PF-0960 USN
;  CURRENT APPLICATION NUMBER: US/10/476,924
;  CURRENT FILING DATE:  2003-11-04
;  PRIOR APPLICATION NUMBER: PCT/US02/14276
;  PRIOR FILING DATE: 2002-05-02
;  PRIOR APPLICATION NUMBER: US 60/288,598
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: US 60/291,776
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: US 60/292,172
;  PRIOR FILING DATE: 2001-05-18
;  PRIOR APPLICATION NUMBER: US 60/293,564
;  PRIOR FILING DATE: 2001-05-25
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PERL Program
; SEQ ID NO 7
;   LENGTH: 802
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 4706628CD1
US-10-476-924-7

  Query Match           59.3%;  Score 1718.5;  DB 16;  Length 802;
  Best Local Similarity   59.2%;  Pred. No. 8.4e-135;
  Matches  316;  Conservative   77;  Mismatches  116;  Indels   25;  Gaps    6;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:||||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:|||||:||||||||:|||||||||||||||||||: |||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:|||||||||: |:|| |: ||||||||::|:||||     | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:|||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              |||||||   | ||||||||||||| || ||| || :| ::| | :||| ||       :
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-----PSD 644

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              :| :  | : ::|  |    |           |    || :   : :     :| : || |
Db        645 KEGKKPKGQSKKQPSGTTKRP---------ISDDDCPSASKVYKASDSAEAIEAFQLTPQ 695

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  |||||  | |||:|||: | : |          || |:|::| |:||||||::|:|| |
Db        696 QQHLIREDCQNQKLWDEVLSHLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 748

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              ||||||| |||:||||||||||:||::| |  |: |||:|:  ||||  ||
Db        749 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYIMIPNEILQTLLDLFFPGYSKGR 802
```

```
RESULT 7
US-10-126-103-113
; Sequence 113, Application US/10126103
; Publication No. US20030224486A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB PATHWA
;  FILE REFERENCE: D0108.np
;  CURRENT APPLICATION NUMBER: US/10/126,103
;  CURRENT FILING DATE:  2002-04-19
;  PRIOR APPLICATION NUMBER: US 60/284,962
;  PRIOR FILING DATE: 2001-04-19
;  PRIOR APPLICATION NUMBER: US 60/286,645
;  PRIOR FILING DATE: 2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/346,986
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113
;   LENGTH: 645
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-126-103-113

  Query Match           44.8%;  Score 1297;  DB 15;  Length 645;
  Best Local Similarity   67.2%;  Pred. No. 1.3e-99;
  Matches  231;  Conservative   48;  Mismatches   61;  Indels    4;  Gaps    3;
```

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||    |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:|||||:||||||||:||||||||||||||||: ||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: ||||||||::|:||||     | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGL 341
              |||||||   | |||||||||||||| || ||| || :| ::| |
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCL 633


RESULT 8
US-10-431-096-113
; Sequence 113, Application US/10431096
; Publication No. US20040086896A1
; GENERAL INFORMATION:
;  APPLICANT: Bristol-Myers Squibb Company
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
;  TITLE OF INVENTION:  PATHWAY
;  FILE REFERENCE: D0108A CIP
;  CURRENT APPLICATION NUMBER: US/10/431,096
;  CURRENT FILING DATE:  2003-05-07
;  PRIOR APPLICATION NUMBER: US 60/284,962
;  PRIOR FILING DATE: 2001-04-19
;  PRIOR APPLICATION NUMBER: US 10/126,103
;  PRIOR FILING DATE: 2002-04-19
;  PRIOR APPLICATION NUMBER: US 60/286,645
;  PRIOR FILING DATE: 2001-04-26
;  PRIOR APPLICATION NUMBER: US 60/346,986
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 307
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 113
;   LENGTH: 645
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-431-096-113

  Query Match             44.8%;  Score 1297;  DB 15;  Length 645;
  Best Local Similarity   67.2%;  Pred. No. 1.3e-99;
  Matches  231;  Conservative   48;  Mismatches   61;  Indels    4;  Gaps    3;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||    |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:|||||:||||||||:||||||||||||||||: ||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: ||||||||::|:||||     | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGL 341
              |||||||   | |||||||||||||| || ||| || :| ::| |
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCL 633


RESULT 9
US-09-867-550-766
; Sequence 766, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 766
;   LENGTH: 198
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (57)
;   OTHER INFORMATION: wherein Xaa may be any one of Arg or Cys or Gly or Ser
US-09-867-550-766

  Query Match             28.4%;  Score 823.5;  DB 9;  Length 198;
  Best Local Similarity   75.8%;  Pred. No. 1.3e-60;
  Matches  150;  Conservative   22;  Mismatches   25;  Indels    1;  Gaps    1;

Qy        122 AKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVHR 181
              ||| ||:: |:||||:||||||||:||||||||||||||||||: |||||||||  |||
Db          1 AKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEEXVHR 60

Qy        182 PHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTN 241
              ||| |||||||||:||||||||: |:|| |: ||||||||::|:||||     | || |||
Db         61 PHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQTLTN 120

Qy        242 TNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVK 301
              |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||||||
Db        121 MNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYKVVK 180

Qy        302 YWPEKGKS-GFLVWRYLL 318
              ||||   | |||||||||
Db        181 YWPEISSSHGFLVWRYLL 198


RESULT 10
US-10-123-568-3
; Sequence 3, Application US/10123568
; Publication No. US20030194713A1
; GENERAL INFORMATION:
;  APPLICANT: Hitoshi, Yasumichi
;  APPLICANT:  Jenkins, Yonchu
;  APPLICANT:  Rigel Pharmaceuticals, Inc.
;  TITLE OF INVENTION: NP95:  Methods of Assaying for Cell Cycle Modulators
;  FILE REFERENCE: 021044-003400US
;  CURRENT APPLICATION NUMBER: US/10/123,568
;  CURRENT FILING DATE:  2002-04-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
```

```
;   LENGTH: 136
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:G1-2635
US-10-123-568-3

  Query Match             25.4%;  Score 736;  DB 14;  Length 136;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-53;
  Matches  136;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        231 AEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEG 290
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEG 60

Qy        291 NRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYL 350
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYL 120

Qy        351 EALANREREKENSKRE 366
              ||||||||||||||||
Db        121 EALANREREKENSKRE 136


RESULT 11
US-09-764-864-1301
; Sequence 1301, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1301
;   LENGTH: 133
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (93)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1301

  Query Match             23.2%;  Score 672;  DB 9;  Length 133;
  Best Local Similarity   99.2%;  Pred. No. 3.8e-48;
  Matches  126;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        405 KKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCIC 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 KKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCIC 66

Qy        465 CQELVFRPITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLF 524
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         67 CQELVFRPITTVCQHNVCKDCLDRSFXAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLF 126

Qy        525 PGYGNGR 531
              |||||||
Db        127 PGYGNGR 133


RESULT 12
US-10-425-114-39797
; Sequence 39797, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 39797
;   LENGTH: 694
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700763470_FLI.pep
US-10-425-114-39797

  Query Match             21.6%;  Score 626.5;  DB 15;  Length 694;
  Best Local Similarity   28.1%;  Pred. No. 2.3e-43;
  Matches  180;  Conservative   66;  Mismatches  161;  Indels  233;  Gaps   21;

Qy         55 ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPEC------------- 101
              || :| |    :::|:|  ||  :|: ||  |  :: :   | ||:|
Db          3 ACMVCKGLPGQEERLLCVTCDTPWHVPCLFAPPPTLSATARWLCPDCSILDSDVPPVPAP 62

Qy        102 ------------RNDAS----------EVVLAG--------------------------- 112
                          ||||          : :| |
Db         63 ARNQLVAAMLAVENDASLTQHDKARKRQELLTGKAPADDDDDDEQENKSSLSDILSRSLN 122

Qy        113 --------ER------------------LRESKKN---------AKMAS----------- 126
                     ||                  :|: |:|         ||| |
Db        123 CSICIQLPERPVTSPCGHSFCLKCFEKWVRQGKRNCAKCRQIIPAKMTSQPRINSALVFA 182

Qy        127 ------ATSSS-------------------------QRDWGKGMACVGRTKECTIVPSN 154
                   |::||                        :|    | |     |    ||::
Db        183 IRMARQASNSSGSGRGGVRSVNHFLHNQDRPDEPFTTERAQRSGRANAASGKIFVTVPTD 242

Qy        155 HYGPI---------PGIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYE 205
              |:|||          |: ||  || |::  : | |   | || |:|: |: |:||:|||
Db        243 HFGPITAENDPLRNQGLLVGESWRDRLECRQWGAHFVPVGGIAGQSDRGAQSVVLSGGYV 302

Qy        206 DDVDHGNFFTYTGSGGRDLSGNKRT-AEQSCDQKLTNTNRALALNCFAPINDQEGAEAKD 264
              || ||| :| |||||:|||||||    | |||    |||| ::|
Db        303 DDEDHGEWFLYTGSGGKDLSGNKRTNKSHSFDQKFEKYNRALQVSCL------------- 349

Qy        265 WRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDE 324
                | ||||||: |  | | ||| | ||||||:: | |   |  || | |||  | |:|
Db        350 --QGYPVRVVRSHK-EKRSSYAPETGVRYDGIYRIEKCWQIAGLQGFKVCRYLFVRCDNE 406

Qy        325 PGPWTKEGK-DRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFASPRTGKG 383
              | ||| :   || : | :              || :|     |  |     |      :
Db        407 PAPWTSDDHGDRPRPLPVI------------RELKKATVIHERTESPSWDF-DEEDSRW 452

Qy        384 KWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRP 443
              |||:     |||       |  |||          :  ||| ::    | |:|:
Db        453 KWKKPPP---PSR-------QKVQNVEPVE--------VARAKSNKEVKKFQLKSIKE-- 492

Qy        444 ASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSFRAQVF-------- 495
                      ::  | |: |:|::  |:|| | || || ||:  |  | |
Db        493 -----------QLQRGFSCMICKEVMVSPVTTPCAHNFCKSCLEGEFAGQAFVKERSKGG 541

Qy        496 ----------SCPACRYDLGRSYA--MQVNQPLQTVLNQL 523
                       ||:|  |:   |   :||:  |:: :  |
Db        542 RTLRSQKNVMKCPSCSIDIS-DYLQNIQVDIDLKSAIESL 580


RESULT 13
US-10-424-599-174378
; Sequence 174378, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
```

```
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174378
;   LENGTH: 709
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_128482C.1.pep
US-10-424-599-174378

  Query Match             21.6%;  Score 626.5;  DB 15;  Length 709;
  Best Local Similarity   28.1%;  Pred. No. 2.3e-43;
  Matches  180;  Conservative   66;  Mismatches  161;  Indels  233;  Gaps   21;

Qy          55 ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPEC---------- 101
               || :| |    :::|:|  ||  :|: ||  |  :: :   | ||:|
Db          18 ACMVCKGLPGQEERLLCVTCDTPWHVPCLFAPPPTLSATARWLCPDCSILDSDVPPVPAP 77

Qy         102 ------------RNDAS----------EVVLAG---------------------------- 112
                         ||||          : :| |
Db          78 ARNQLVAAMLAVENDASLTQHDKARKRQELLTGKAPADDDDDDEQENKSSLSDILSRSLN 137

Qy         113 --------ER-----------------LRESKKN---------AKMAS------------ 126
                     ||                  :|: |:|         ||| |
Db         138 CSICIQLPERPVTSPCGHSFCLKCFEKWVRQGKRNCAKCRQIIPAKMTSQPRINSALVFA 197

Qy         127 ------ATSSS--------------------------QRDWGKGMACVGRTKECTIVPSN 154
                    |::||                        :|    | |      |    ||::
Db         198 IRMARQASNSSGSGRGGVRSVNHFLHNQDRPDEPFTTERAQRSGRANAASGKIFVTVPTD 257

Qy         155 HYGPI---------PGIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYE 205
               |:|||          |: ||  || |::  : | |   | || |:|: |: |:||:|||
Db         258 HFGPITAENDPLRNQGLLVGESWRDRLECRQWGAHFVPVGGIAGQSDRGAQSVVLSGGYV 317

Qy         206 DDVDHGNFFTYTGSGGRDLSGNKRT-AEQSCDQKLTNTNRALALNCFAPINDQEGAEAKD 264
               || ||| :| |||||||:||||||||    | |||    |||| ::|
Db         318 DDEDHGEWFLYTGSGGKDLSGNKRTNKSHSFDQKFEKYNRALQVSCL------------- 364

Qy         265 WRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDE 324
                 | |||||: |  | | |||  | ||||||:: | |    |  || | |||  | |:|
Db         365 --QGYPVRVVRSHK-EKRSSYAPETGVRYDGIYRIEKCWQIAGLQGFKVCRYLFVRCDNE 421

Qy         325 PGPWTKEGK-DRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFASPRTGKG 383
               | ||| :   || : | :              || :|    | |      |       :
Db         422 PAPWTSDDHGDRPRPLPVI------------RELKKATVIHERTESPSWDF-DEEDSRW 467

Qy         384 KWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRP 443
               |||:      |||         |   |||          :   ||| ::    | |:|:
Db         468 KWKKPPP---PSR-------QKVQNVEPVE--------VARAKSNKEVKKFQLKSIKE-- 507

Qy         444 ASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSFRAQVF-------- 495
                    :::  | |: |:|::  |:|| | || || ||:  |  | |
Db         508 -----------QLQRGFSCMICKEVMVSPVTTPCAHNFCKSCLEGEFAGQAFVKERSKGG 556

Qy         496 ----------SCPACRYDLGRSYA--MQVNQPLQTVLNQL 523
                        ||:|  |:   |   :||:  |:: :  |
Db         557 RTLRSQKNVMKCPSCSIDIS-DYLQNIQVDIDLKSAIESL 595


RESULT 14
US-10-739-930-6375
; Sequence 6375, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
;  APPLICANT: Kovalic, David K.
;  TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
;  TITLE OF INVENTION:  PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
;  FILE REFERENCE: 38-21(53377)B
;  CURRENT APPLICATION NUMBER: US/10/739,930
;  CURRENT FILING DATE:  2003-12-18
;  NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 6375
;   LENGTH: 645
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: Clone ID: ARATH-23APR03-C352_1.p
US-10-739-930-6375

  Query Match             21.0%;  Score 608;  DB 16;  Length 645;
  Best Local Similarity   26.7%;  Pred. No. 7.3e-42;
  Matches  173;  Conservative   69;  Mismatches  161;  Indels  246;  Gaps   19;

Qy          56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG--- 112
               |  |     |:: | |  |   :|: ||  |  :: |  :|:||:|  :   : ::|
Db          15 CMRCKSNPPPEESLTCGTCVTPWHVSCLSSPPKTLASTLQWHCPDCSGEIDPLPVSGGAT 74

Qy         113 ----------------------------------------------ERLRESK-------- 119
                                                            |  |: |
Db          75 GFESAGSDLVAAIRAIEADESLSTEEKAKMRQRLLSGKGVEEDDEEEKRKKKGKGKNPNL 134

Qy         120 ------------------------------------------------------------ 119

Db         135 DVLSALGDNLMCSFCMQLPERPVTKPCGHNACLKCFEKWMGQGKRTCGKCRSIIPEKMAK 194

Qy         120 ---------------KNAKMASATSS------SQRD-------------WGKGMACVGRT 145
                             | :| |:||:|      | :|              ||  |  |
Db         195 NPRINSSLVAAIRLAKVSKSAAATTSKVFHFISNQDRPDKAFTTERAKKTGKANAASG-- 252

Qy         146 KECTIVPSNHYGPIP---------GIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSY 196
               |    :| :|:||||         |: ||  |  |::  : | | |||||| |:|  |:
Db         253 KIYVTIPPDHFGPIPAENDPVRNQGLLVGESWEDRLECRQWGAHFPHVAGIAGQSTYGAQ 312

Qy         197 SLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRT-AEQSCDQKLTNTNRALALNCFAPIN 255
               |: |:|||:|| ||| :| ||||||||||||||  ||| |||   :| || |:|
Db         313 SVALSGGYKDDEDHGEWFLYTGSGGRDLSGNKRTNKEQSFDQKFEKSNAALKLSC----- 367

Qy         256 DQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSG-FLVW 314
                       : | ||||||: |  | | ||| || ||||:|:: | | : |  | | |
Db         368 ----------KLGYPVRVVRSHK-EKRSAYAPEEGVRYDGVYRIEKCWRKVGVQGSFKVC 416

Qy         315 RYLLRRDDDEPGPWTK-EGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEG 373
               |||  | |:|| |||  |  || : :       ||  |     :    ||     : :|
Db         417 RYLFVRCDNEPAPWTSDENGDRPRPI---PNIPE--LNMATDLFERKETPSWDFDE---- 467

Qy         374 GFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWN 433
                     |:| ||              |   |||:          :  |   |:: | :
Db         468 -------GEGCWKWM----------KPPPASKKS----------VNVLAPEERKNLR--- 497

Qy         434 EVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSF--- 490
                       :    :      ::: : |:|  ||:::  |:|| | || || ||:  |
Db         498 --------KAIKAAHSNTMRARLLKEFKCQICQQVLTLPVTTPCAHNFCKACLEAKFAGK 549

Qy         491 ---------------RAQVFSCPACRYDLGRSYA-MQVNQPLQTVLNQL 523
                            |  | :|| |  |:       |||: :  |: :|
Db         550 TLVRERSTGGRTLRSRKNVLNCPCCPTDISDFLQNPQVNREVAEVIEKL 598


RESULT 15
US-10-225-066A-522
; Sequence 522, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
;  APPLICANT: Mendel Biotechnology, Inc.
;  APPLICANT:  RATCLIFFE, Oliver
;  APPLICANT:  RIECHMANN, Jose Luis
;  APPLICANT:  ADAM, Luc J
;  APPLICANT:  DUBELL, Arnold T
;  APPLICANT:  HEARD, Jacqueline E
;  APPLICANT:  PILGRIM, Marsha L
;  APPLICANT:  JIANG, Cai-Zhong
```

```
;  APPLICANT:  REUBER, T. Lynne
;  APPLICANT:  CREELMAN, Robert A
;  APPLICANT:  PINEDA, Omaira
;  APPLICANT:  YU, Guo-Liang
;  APPLICANT:  BROUN, Pierre E
;  TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
;  FILE REFERENCE: MBI0036-2 US
;  CURRENT APPLICATION NUMBER: US/10/225,066A
;  CURRENT FILING DATE:  2002-08-09
;  PRIOR APPLICATION NUMBER: 09/837,444
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 1122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 522
;   LENGTH: 617
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-225-066A-522

  Query Match             20.3%;  Score 588.5;  DB 15;  Length 617;
  Best Local Similarity   27.6%;  Pred. No. 3e-40;
  Matches  174;  Conservative   58;  Mismatches  161;  Indels  237;  Gaps   22;

Qy          56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPEC---------- 101
               |  |      :: | |  |   :|: || |  |   |  || ||:|
Db          15 CMRCQVNPPSEETLTCGTCVTPWHVPCLLPE-SLASSTGEWECPDCSGVVVPSAAPGTGN 73

Qy         102 -RNDASEVVL--------AGERLRESKKNAKMASATSSSQRD---------------- 134
                | ::|  ||        | | | |::|  |     |    |
Db          74 ARPESSGSVLVAAIRAIQADETLTEAEKAKKRQKLMSGGGDDGVDEEEKKKLEIFCSICI 133

Qy         135 ----------------------WGKG------MAC----------------------VG 143
                                    |  |      | |                      :
Db         134 QLPERPITTPCGHNFCLKCFEKWAVGQGKLTCMICRSKIPRHVAKNPRINLALVSAIRLA 193

Qy         144 RTKECTI-----------------------------------------VPSNHYGPIP- 160
               :|::                                         :| :|:||||
Db         194 NVTKCSVEATAAKVHHIIRNQDRPEKAFTTERAVKTGKANAASGKFFVTIPRDHFGPIPA 253

Qy         161 --------GIPVGTMWRFRVQVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGN 212
                     |: ||  |  | :  : | | ||:||| |:|  |: |: |:|||:|| |||
Db         254 ENDVTRKQGVLVGESWEDRQECRQWGAHFPHIAGIAGQSAVGAQSVALSGGYDDDEDHGE 313

Qy         213 FFTYTGSGGRDLSGNKR-TAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPV 271
               :| ||||||||||||||   :|| ||   | | :| |:|                : | ||
Db         314 WFLYTGSGGRDLSGNKRINKKQSSDQAFKNMNESLRLSC---------------KMGYPV 358

Qy         272 RVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSG-FLVWRYLLRRDDDEPGPWTK 330
               ||||: |  | | |||||| ||||:|:: | |   |  | | | |||  | |:|| |||
Db         359 RVVRSWK-EKRSAYAPAEGVRYDGVYRIEKCWSNVGVQGSFKVCRYLFVRCDNEPAPWTS 417

Qy         331 -EGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKS 389
                |  || : |      ||  ||  |:    ||:   : :| :          |: || :
Db         418 DEHGDRPRPL---PNVPE--LETAADLFVRKESPSWDFDEAE---------GRWKWMK-- 461

Qy         390 AGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPF 449
                        ||           | |  |        | | | :  | : | |
Db         462 ---------SP----------PVSRMALDP---EERKKNKRAKNTMKARLL--------- 490

Qy         450 QLFLSKVEETFQCICCQELVFRPITTVCQHNVCKDCLDRSF-----------------R 491
                     : | |  |:|::  |:|| | || || ||:  |                 :
Db         491 --------KEFSCQICREVLSLPVTTPCAHNFCKACLEAKFAGITQLRERSNGGRKLRAK 542

Qy         492 AQVFSCPACRYDLGRSYA-MQVNQPLQTVL 520
                 : :|| |  ||       |||: :  ::
```

```
Db         543 KNIMTCPCCTTDLSEFLQNPQVNREMMEII 572

Search completed: July 25, 2005, 17:11:06
Job time : 160 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:56:22 ; Search time 43 Seconds
(without alignments)
921.829 Million cell updates/sec

Title: US-10-019-071-2_COPY_263_793
Perfect score: 2896
Sequence: 1 DDSLNDCRIIFVDEVFKIER.....VNQPLQTVLNQLFPGYGNGR 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	186.5	6.4	1400	4	US-09-764-176-7	Sequence 7, Appli
2	177	6.1	4019	4	US-09-854-133-425	Sequence 425, App
3	157.5	5.4	2079	4	US-09-949-016-8301	Sequence 8301, Ap
4	152.5	5.3	1935	4	US-09-949-016-10970	Sequence 10970, A
5	152	5.2	1674	4	US-09-418-710-1	Sequence 1, Appli
6	152	5.2	1674	4	US-09-839-479-1	Sequence 1, Appli
7	149	5.1	1673	4	US-09-418-710-70	Sequence 70, Appl
8	149	5.1	1673	4	US-09-839-479-69	Sequence 69, Appl
9	147	5.1	1527	4	US-09-418-710-27	Sequence 27, Appl
10	147	5.1	1527	4	US-09-839-479-27	Sequence 27, Appl
11	147	5.1	1531	4	US-09-418-710-29	Sequence 29, Appl
12	147	5.1	1531	4	US-09-839-479-29	Sequence 29, Appl
13	147	5.1	1540	4	US-09-949-016-7037	Sequence 7037, Ap
14	147	5.1	1722	4	US-09-538-092-1033	Sequence 1033, Ap
15	146.5	5.1	2289	3	US-09-051-019-2	Sequence 2, Appli
16	145.5	5.0	800	4	US-09-270-767-45282	Sequence 45282, A
17	143	4.9	1525	4	US-09-418-710-69	Sequence 69, Appl
18	143	4.9	1525	4	US-09-839-479-68	Sequence 68, Appl
19	143	4.9	1912	3	US-08-913-832A-2	Sequence 2, Appli
20	143	4.9	1912	4	US-09-249-181A-2	Sequence 2, Appli
21	143	4.9	1912	4	US-09-158-707-2	Sequence 2, Appli
22	139	4.8	1969	4	US-09-418-710-72	Sequence 72, Appl
23	139	4.8	1969	4	US-09-839-479-71	Sequence 71, Appl
24	139	4.8	1972	4	US-09-418-710-21	Sequence 21, Appl
25	139	4.8	1972	4	US-09-839-479-21	Sequence 21, Appl
26	138.5	4.8	45	4	US-09-839-479-44	Sequence 44, Appl
27	138.5	4.8	351	4	US-07-945-295-2	Sequence 2, Appli
28	138.5	4.8	351	5	PCT-US91-06418-1	Sequence 1, Appli
29	138.5	4.8	504	4	US-09-270-767-45920	Sequence 45920, A
30	138	4.8	386	4	US-09-764-176-3	Sequence 3, Appli
31	137.5	4.7	371	3	US-09-233-342A-5	Sequence 5, Appli
32	135.5	4.7	391	4	US-09-538-092-1369	Sequence 1369, Ap
33	135.5	4.7	391	4	US-09-949-016-6711	Sequence 6711, Ap
34	135.5	4.7	391	4	US-09-949-016-7788	Sequence 7788, Ap
35	135.5	4.7	405	2	US-08-881-857-2	Sequence 2, Appli
36	135.5	4.7	405	3	US-09-233-342A-2	Sequence 2, Appli
37	134.5	4.6	45	4	US-09-418-710-43	Sequence 43, Appl
38	134.5	4.6	45	4	US-09-418-710-56	Sequence 56, Appl
39	134.5	4.6	45	4	US-09-839-479-42	Sequence 42, Appl
40	134.5	4.6	45	4	US-09-839-479-55	Sequence 55, Appl
41	131.5	4.5	45	4	US-09-418-710-45	Sequence 45, Appl
42	131	4.5	44	4	US-09-418-710-58	Sequence 58, Appl
43	131	4.5	44	4	US-09-839-479-57	Sequence 57, Appl
44	127	4.4	594	4	US-09-949-016-9261	Sequence 9261, Ap
45	125.5	4.3	835	4	US-09-949-016-9685	Sequence 9685, Ap

ALIGNMENTS

RESULT 1
US-09-764-176-7
; Sequence 7, Application US/09764176
; Patent No. 6809189
; GENERAL INFORMATION:
; APPLICANT: NOTEBORN, Mathieu Hubertus Maria
; APPLICANT: DANEN-VAN OORSCHOT, Astrid Adriana Anna Maria
; APPLICANT: ROHN, Jennifer Leigh
; APPLICANT: WEISS, Bertram
; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN
; FILE REFERENCE: 4725US
; CURRENT APPLICATION NUMBER: US/09/764,176
; CURRENT FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 1400
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Description of Sequence: Amino acid sequence deduced from the nu
; OTHER INFORMATION: leic acid seqeuence of AAP-
US-09-764-176-7

Query Match 6.4%; Score 186.5; DB 4; Length 1400;
Best Local Similarity 19.5%; Pred. No. 3.5e-08;
Matches 95; Conservative 64; Mismatches 168; Indels 159; Gaps 19;

```
Qy          56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRND---------- 104
               |  ||    |:  |:|| ||  :|  || |||  :| : ||:|| |::
Db         853 CKKCGLPNHPELILLCDSCDSGYHTACLRPPLMIIP-DGEWFCPPCQHKLLCEKLEEQLQ 911

Qy         105 --------------ASEVVLAG--------------ERLRESKKNAKMASATSSSQRD 134
                           :|  |                |   | ||::| : |    :|
Db         912 DLDVALKKKERAERRKERLVYVGISIENIIPPQEPDFSEDQEEKKKDSKKSKANLLERRS 971

Qy         135 WGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRV------QVSESGVHRPHVAGIH 188
                  ||::|                :||       :  |  :       |:
Db         972 --------TRTRKC----------------ISYRFDEFDEAIDEAIEEDDIKEADGGGV- 1005

Qy         189 GRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLS--------GNKRTAEQSCDQKLT 240
               ||  | |                    | ||  |:|:|          |||  :::   :
Db        1006 GRGKDIS------------------TITGHRGKDISTILDEERKENKR-PQRAAAARRK 1045

Qy         241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIY--- 297
                |   |:  : ::::|  :       |      | : :    |:  |   :  |  :
Db        1046 KRRRLNDLDSDSNLDEEESEDEFKISDGSQDEFVVSDENPDESEEDPPSNDDSDTDFCSR 1105

Qy         298 -------------KVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTM 343
```

```
Db       1106 RLRRHPSRPMRQSRRLRRKTPKKKYS----------DDDEEE--ESEENSRDSESDFSD 1152

Qy        344 QYPEGYLEALANREREKENSKRE----EEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGS 399

Db       1153 DFSDDFVE--TRRRRSRRNQKRQINYKEDSESDGSQKSLRRGK----------EIRRVH 1199

Qy        400 PRRTSKKTKVEPY-SLTAQQSSLIREDKSNAKL-------WNEVLASLKDRPASGSPFQL 451

Db       1200 KRRLSSSESEESYLSKNSEDDELAKESKRSVRKRGRSTDEYSE--ADEEEEEEEGKPSRK 1257

Qy        452 FLSKVE 457

Db       1258 RLHRIE 1263


RESULT 2
US-09-854-133-425
; Sequence 425, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
;  APPLICANT: Lodes, Michael J.
;  APPLICANT:  Mohamath, Raodoh
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Benson, Darin R.
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  THE THERAPY AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.475C10
;  CURRENT APPLICATION NUMBER: US/09/854,133
;  CURRENT FILING DATE:  2001-05-11
;  NUMBER OF SEQ ID NOS: 735
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 425
;   LENGTH: 4019
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-854-133-425

  Query Match             6.1%;  Score 177;  DB 4;  Length 4019;
  Best Local Similarity   21.6%;  Pred. No. 1.4e-06;
  Matches  109;  Conservative   58;  Mismatches  150;  Indels  188;  Gaps   25;

Qy          51 CRVC-ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEW---YCPECRN-DA 105

Db          59 CLECTVCEACGKATDPGRLLLCDDCDISYHTYCLDPPLQTVP-KGGWKCKWCVWCRHCGA 117

Qy         106 SEVVLAGERLRESKKNAKMASATS-----SSQRDWGKGMACVGRTKECTIVPSNHYGPIP 160

Db         118 TSAGLRCEWQNNYTQCAPCASLSSCPVCYRNYREEDLILQC----RQC------------ 161

Qy         161 GIPVGTMWRFRV----------QVSESG----VHRPHVAGIHGRSNDGSYSLVLAG--G 203

Db         162 -----DRWMHAVCQNLNTEEEVENVADIGFDCSMCRPYMPASNVPSSDCCESSLVAQIVT 216

Qy         204 YEDDVDHGNFFTYTG-----SGGRDLSG------NKRTAEQSCDQKLTNTNRALALNCFA 252

Db         217 KVKELDPPKTYTQDGVCLTESGMTQLQSLTVTVPRRKRSKPKLKLKIINQNSVAVLQTPP 276

Qy         253 PINDQEG--AEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSG 310

Db         277 DIQSEHSRDGEMDDSREGE----LMDCDGKSESS------------------PEREAV- 312

Qy         311 FLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREKENSKREEEEQ 370

Db         313 ----------DDETKG---VEGTDGVKK-----------------------RKRKPYRP 335

Qy         371 QEGGF---ASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSSLIREDKS 427

Db         336 GIGGFMVRQRSRTGQGKTKR------------------------------SVIRKDSS 363

Qy         428 NA---------KLWNEVLA-SLKDRPASGSP--------FQLFLSKVEETFQCICCQELV 469

Db         364 GSISEQLPCRDDGWSEQLPDTLVDESVSVTESTEKIKKRYRKRKNKLEETFPAY-LQEAF 422

Qy         470 FRPITTVCQHNVCKDCLDRSFRAQV 494

Db         423 FG----------KDLLDTSRQSKI 436


RESULT 3
US-09-949-016-8301
; Sequence 8301, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8301
;   LENGTH: 2079
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-8301

  Query Match             5.4%;  Score 157.5;  DB 4;  Length 2079;
  Best Local Similarity   32.0%;  Pred. No. 3.6e-05;
  Matches   48;  Conservative   14;  Mismatches   49;  Indels   39;  Gaps    9;

Qy          38 PSC-KHCKD---DVNRL------CRVC-ACHLCGGRQDPDKQLMCDECDMAFHIYCLDPP 86

Db         252 PSCLKFCPELTTNVKALRWQCIECKTCSACRVQG--RNADNMLFCDSCDRGFHMECCDPP 309

Qy          87 LSSVPSEDEWYCPECRNDASEVVLAGERLRESKK----------NAKMASATSSSQR 133

Db         310 LSRMP-KGMWICQVCRPKKKGRKLLHEKAAQIKRRYAKPIGRPKNKLKQRLLSVTS---- 364

Qy         134 DWGKGMACVGR--------TKECTIVPSNH 155

Db         365 DEGSMNAFTGRGSPGRGQKTKVCTTPSSGH 394


RESULT 4
US-09-949-016-10970
; Sequence 10970, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10970
;   LENGTH: 1935
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10970
```

```
  Query Match             5.3%;  Score 152.5;  DB 4;  Length 1935;
  Best Local Similarity   20.5%;  Pred. No. 9.6e-05;
  Matches   91;  Conservative   47;  Mismatches  120;  Indels  185;  Gaps   22;

Qy         39 SCKHC----------------------------KDDVNRLCRVCACHLCGGRQDPDKQLM 70
              || ||                            :||    ||||        :|  : |
Db        428 SCPHCEKEGVQWEAKEEEEEYEEEGEEEGEKEEEDDHMEYCRVC--------KDGGELLC 479

Qy         71 CDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG--ERLRESKKNAKMASAT 128
              || |  ::||:||:|||  :|:  || || |    :  || |  :::  :
Db        480 CDACISSYHIHCLNPPLPDIPN-GEWLCPRC----TCPVLKGRVQKILHWR--------- 525

Qy        129 SSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVHRPHVAGIH 188
                    ||:    |               | |            | ::     |    :
Db        526 ------WGEPPVAV--------------PAP-----------QQADGNPDVPPPRPLQ 552

Qy        189 GRSND----------------------GSYSLVLAGGYE--DDVDHGNFFTYTGSGGRDL 224
              |||                        : ||:    |:  :|:|      | |||  |
Db        553 GRSEREFFVKWVGLSYWHCSWAKELQLEIFHLVMYRNYQRKNDMDEPPPLDY-GSGEDDG 611

Qy        225 SGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKD--WRSG-KPVRVVRNVKGGK 281
              :||                        : |   || ::  :| | ||
Db        612 KSDKR----------------------KVKDPHYAEMEEKYYRFGIKP----------- 637

Qy        282 NSKYAPAEGNRYDGIYKVVKY-WPEKGKSGFLV-WRYLLRRDDDEPGPWTKEGKDRIKKL 339
                        :  ::::: : .  :||   :|| || |      :   | ::
Db        638 ----------EWMTVHRIINHSVDKKGNYHYLVKWRDL----PYDQSTWEED-------- 675

Qy        340 GLTMQYPEGYLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGS 399
                |  ||          |  |::  |  |       | ||  | | |     | ||   |
Db        676 --EMNIPE--------YEEHKQSYWRHRELIMGEDPAQPRKYKKKKKELQGDGPPS---S 722

Qy        400 PRR-TSKKTKVEPYSLTAQQSSL 421
              |    : | : :|  :||   :|
Db        723 PTNDPTVKYETQPRFITATGGTL 745


RESULT 5
US-09-418-710-1
; Sequence 1, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
;  APPLICANT:  Jones, Michael H.
;  TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE:  06501-042001
;  CURRENT APPLICATION NUMBER: US/09/418,710
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER:  PCT/JP98/01783
;  PRIOR FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER:  JP 9/310027
;  PRIOR FILING DATE:  1997-10-24
;  PRIOR APPLICATION NUMBER:  JP 9/116570
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-418-710-1

  Query Match             5.2%;  Score 152;  DB 4;  Length 1674;
  Best Local Similarity   17.9%;  Pred. No. 8.6e-05;
  Matches   87;  Conservative   61;  Mismatches  144;  Indels  194;  Gaps   20;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : | :  ::|| ||   | ||: | | :|| | :|:|||||
Db       1269 CKICRKKGDAENMVLCDGCDRGHHTYCVRPKLKTVP-EGDWFCPECR------------ 1314

Qy        116 RESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVS 175
                 |:  :  |                                              ||
Db       1315 --PKQRCRRLS-------------------------------------------FR---- 1325

Qy        176 ESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSC 235
                 || :        |   |:   || :|:||           | :  |     |
Db       1326 ----QRPSLES----DEDVEDSM---GGEDDEVD----------GDEEEGQSEEEEYEV 1363

Qy        236 DQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKP-VRVVRNVKGGKNSKYAPAEGNRYD 294
              :|  ::              ||  |    : |:| ||:    :|  :| :: :
Db       1364 EQDEDDS-------------QEEEEVSLPKRGRPQVRLPVKTRGKLSSSFSSRGQQQEP 1409

Qy        295 GIYKVVKYWPEK-------------GKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLG 340
              | |      | :              |:|       ||: :  |   ||    ::
Db       1410 GRY------PSRSQQSTPKTTVSSKTGRS--------LRKINSAPPTETKS----LRIAS 1451

Qy        341 LTMQYPEG------YLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGP 394
               : ::  |      ::| |: | : :          :    | : |    |   :|
Db       1452 RSTRHSHGPLQADVFVELLSPRRKRRGRKSANNTPENSPNFPNFRVIATKSSEQSRSVNI 1511

Qy        395 S------RAGSPRRTSKKTKVEPYSLTAQQSS-----------------LIREDKSNA 429
              :      : | ||  |:   | ||  :|  : |                    :| | |
Db       1512 ASKLSLQESESKRRCRKRQSPEPSPVTLGRRSSGRQGGVHELSAFEQLVVELVRHDDS-- 1569

Qy        430 KLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPI--------TTVCQHNV 481
               |              ||   :||:    |     ::: :||           |:: :
Db       1570 --W--------------PFLKLVSKI----QVPDYYDIIKKPIALNIIREKVNKCEYKL 1608

Qy        482 CKDCLD 487
                : :|
Db       1609 ASEFID 1614


RESULT 6
US-09-839-479-1
; Sequence 1, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
;  APPLICANT: Jones, Michael H.
;  TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE: 06501-042002
;  CURRENT APPLICATION NUMBER: US/09/839,479
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 09/418,710
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: PCT/JP98/01783
;  PRIOR FILING DATE: 1998-04-17
;  PRIOR APPLICATION NUMBER: JP 9/310027
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: JP 9/116570
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 1674
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-839-479-1

  Query Match             5.2%;  Score 152;  DB 4;  Length 1674;
  Best Local Similarity   17.9%;  Pred. No. 8.6e-05;
  Matches   87;  Conservative   61;  Mismatches  144;  Indels  194;  Gaps   20;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : | :  ::|| ||   | ||: | | :|| | :|:|||||
Db       1269 CKICRKKGDAENMVLCDGCDRGHHTYCVRPKLKTVP-EGDWFCPECR------------ 1314

Qy        116 RESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVS 175
                 |:  :  |                                              ||
Db       1315 --PKQRCRRLS-------------------------------------------FR---- 1325

Qy        176 ESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSC 235
                 || :        |   |:   || :|:||           | :  |     |
Db       1326 ----QRPSLES----DEDVEDSM---GGEDDEVD----------GDEEEGQSEEEEYEV 1363

Qy        236 DQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKP-VRVVRNVKGGKNSKYAPAEGNRYD 294
```

```
Db       1364 EQDEDDS-------------QEEEEVSLPKRGRPQVRLPVKTRGKLSSSFSSRGQQQEP 1409

Qy        295 GIYKVVKYWPEK-------------GKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLG 340
Db       1410 GRY------PSRSQQSTPKTTVSSKTGRS--------LRKINSAPPTETKS----LRIAS 1451

Qy        341 LTMQYPEG------YLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGP 394
Db       1452 RSTRHSHGPLQADVFVELLSPRRKRRGRKSANNTPENSPNFPNFRVIATKSSEQSRSVNI 1511

Qy        395 S------RAGSPRRTSKKTKVEPYSLTAQQSS------------------LIREDKSNA 429
Db       1512 ASKLSLQESESKRRCRKRQSPEPSPVTLGRRSSGRQGGVHELSAFEQLVVELVRHDDS-- 1569

Qy        430 KLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPI--------TTVCQHNV 481
Db       1570 --W---------------PFLKLVSKI----QVPDYYDIIKKPIALNIIREKVNKCEYKL 1608

Qy        482 CKDCLD 487
Db       1609 ASEFID 1614


RESULT 7
US-09-418-710-70
; Sequence 70, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
;  APPLICANT:  Jones, Michael H.
;  TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE:  06501-042001
;  CURRENT APPLICATION NUMBER: US/09/418,710
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER:  PCT/JP98/01783
;  PRIOR FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER:  JP 9/310027
;  PRIOR FILING DATE:  1997-10-24
;  PRIOR APPLICATION NUMBER:  JP 9/116570
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 70
;   LENGTH: 1673
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-418-710-70

  Query Match             5.1%;  Score 149;  DB 4;  Length 1673;
  Best Local Similarity   17.9%;  Pred. No. 0.00017;
  Matches   87;  Conservative   60;  Mismatches  145;  Indels  194;  Gaps   20;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
Db       1268 CKICRKKGDAENMVLCDGCDRGHHTYCVRPKLKIVP-EGDWFCPECR------------- 1313

Qy        116 RESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVS 175
Db       1314 --PKQRCRRLS---------------------------------------------FR---- 1324

Qy        176 ESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSC 235
Db       1325 ----QRPSLES----DEDVEDSM---GGEDDEVD-----------GDEEEGQSEEEEYEV 1362

Qy        236 DQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKP-VRVVRNVKGGKNSKYAPAEGNRYD 294
Db       1363 EQDEDDS-------------QEEEEVSLPKRGRPQVRLPVKTRGKLSSSFSSRGQQQEP 1408

Qy        295 GIYKVVKYWPEK-------------GKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLG 340
Db       1409 GRY------PSRSQQSTPKTTVSSKTGRS--------LRKINSAPPTETKS----LRIAS 1450

Qy        341 LTMQYPEG------YLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGP 394
Db       1451 RSTRHSHGPLQADVFVELLSPRRKRRGRKSANNTPENSPNFPNFRVIATKSSEQSRSVNI 1510

Qy        395 S------RAGSPRRTSKKTKVEPYSLTAQQSS------------------LIREDKSNA 429
Db       1511 ASKLSLQESESKRRCRKRQSPEPSPVTLGRRSSGRQGGVHELSAFEQLVVELVRHDDS-- 1568

Qy        430 KLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPI--------TTVCQHNV 481
Db       1569 --W---------------PFLKLVSKI----QVPDYYDIIKKPIALNIIREKVNKCEYKL 1607

Qy        482 CKDCLD 487
Db       1608 ASEFID 1613


RESULT 8
US-09-839-479-69
; Sequence 69, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
;  APPLICANT: Jones, Michael H.
;  TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE: 06501-042002
;  CURRENT APPLICATION NUMBER: US/09/839,479
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 09/418,710
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: PCT/JP98/01783
;  PRIOR FILING DATE: 1998-04-17
;  PRIOR APPLICATION NUMBER: JP 9/310027
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: JP 9/116570
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69
;   LENGTH: 1673
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-839-479-69

  Query Match             5.1%;  Score 149;  DB 4;  Length 1673;
  Best Local Similarity   17.9%;  Pred. No. 0.00017;
  Matches   87;  Conservative   60;  Mismatches  145;  Indels  194;  Gaps   20;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
Db       1268 CKICRKKGDAENMVLCDGCDRGHHTYCVRPKLKIVP-EGDWFCPECR------------- 1313

Qy        116 RESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVS 175
Db       1314 --PKQRCRRLS---------------------------------------------FR---- 1324

Qy        176 ESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSC 235
Db       1325 ----QRPSLES----DEDVEDSM---GGEDDEVD-----------GDEEEGQSEEEEYEV 1362

Qy        236 DQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKP-VRVVRNVKGGKNSKYAPAEGNRYD 294
Db       1363 EQDEDDS-------------QEEEEVSLPKRGRPQVRLPVKTRGKLSSSFSSRGQQQEP 1408

Qy        295 GIYKVVKYWPEK-------------GKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLG 340
Db       1409 GRY------PSRSQQSTPKTTVSSKTGRS--------LRKINSAPPTETKS----LRIAS 1450

Qy        341 LTMQYPEG------YLEALANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGP 394
Db       1451 RSTRHSHGPLQADVFVELLSPRRKRRGRKSANNTPENSPNFPNFRVIATKSSEQSRSVNI 1510

Qy        395 S------RAGSPRRTSKKTKVEPYSLTAQQSS------------------LIREDKSNA 429
Db       1511 ASKLSLQESESKRRCRKRQSPEPSPVTLGRRSSGRQGGVHELSAFEQLVVELVRHDDS-- 1568
```

```
Qy        430 KLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPI--------TTVCQHNV 481
              |                ||   :||:    |      ::: :||           |:: :
Db       1569 --W---------------PFLKLVSKI----QVPDYYDIIKKPIALNIIREKVNKCEYKL 1607

Qy        482 CKDCLD 487
               : :|
Db       1608 ASEFID 1613


RESULT 9
US-09-418-710-27
; Sequence 27, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
;  APPLICANT:  Jones, Michael H.
;  TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE:  06501-042001
;  CURRENT APPLICATION NUMBER: US/09/418,710
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER:  PCT/JP98/01783
;  PRIOR FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER:  JP 9/310027
;  PRIOR FILING DATE:  1997-10-24
;  PRIOR APPLICATION NUMBER:  JP 9/116570
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 1527
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-418-710-27

  Query Match             5.1%;  Score 147;  DB 4;  Length 1527;
  Best Local Similarity   36.7%;  Pred. No. 0.00022;
  Matches   29;  Conservative   13;  Mismatches   29;  Indels    8;  Gaps    2;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : : || ::|||: |||::|| | |  || : || || |:        |  |
Db       1187 CKVCPKKGEDDKLILCDECNKAFHLFCLRPALYEVP-DGEWQCPACQP-------ATARR 1238

Qy        116 RESKKNAKMASATSSSQRD 134
                :|    ||:  |: |
Db       1239 NSRGRNYTEESASEDSEDD 1257


RESULT 10
US-09-839-479-27
; Sequence 27, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
;  APPLICANT: Jones, Michael H.
;  TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE: 06501-042002
;  CURRENT APPLICATION NUMBER: US/09/839,479
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 09/418,710
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: PCT/JP98/01783
;  PRIOR FILING DATE: 1998-04-17
;  PRIOR APPLICATION NUMBER: JP 9/310027
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: JP 9/116570
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 27
;   LENGTH: 1527
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-839-479-27

  Query Match             5.1%;  Score 147;  DB 4;  Length 1527;
  Best Local Similarity   36.7%;  Pred. No. 0.00022;
  Matches   29;  Conservative   13;  Mismatches   29;  Indels    8;  Gaps    2;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : : || ::|||: |||::|| | |  || : || || |:        |  |
Db       1187 CKVCPKKGEDDKLILCDECNKAFHLFCLRPALYEVP-DGEWQCPACQP-------ATARR 1238

Qy        116 RESKKNAKMASATSSSQRD 134
                :|    ||:  |: |
Db       1239 NSRGRNYTEESASEDSEDD 1257


RESULT 11
US-09-418-710-29
; Sequence 29, Application US/09418710
; Patent No. 6596482
; GENERAL INFORMATION:
;  APPLICANT:  Jones, Michael H.
;  TITLE OF INVENTION:  TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE:  06501-042001
;  CURRENT APPLICATION NUMBER: US/09/418,710
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER:  PCT/JP98/01783
;  PRIOR FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER:  JP 9/310027
;  PRIOR FILING DATE:  1997-10-24
;  PRIOR APPLICATION NUMBER:  JP 9/116570
;  PRIOR FILING DATE:  1997-04-18
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 1531
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-418-710-29

  Query Match             5.1%;  Score 147;  DB 4;  Length 1531;
  Best Local Similarity   36.7%;  Pred. No. 0.00022;
  Matches   29;  Conservative   13;  Mismatches   29;  Indels    8;  Gaps    2;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : : || ::|||: |||::|| | |  || : || || |:        |  |
Db       1191 CKVCPKKGEDDKLILCDECNKAFHLFCLRPALYEVP-DGEWQCPACQP-------ATARR 1242

Qy        116 RESKKNAKMASATSSSQRD 134
                :|    ||:  |: |
Db       1243 NSRGRNYTEESASEDSEDD 1261


RESULT 12
US-09-839-479-29
; Sequence 29, Application US/09839479
; Patent No. 6727222
; GENERAL INFORMATION:
;  APPLICANT: Jones, Michael H.
;  TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
;  FILE REFERENCE: 06501-042002
;  CURRENT APPLICATION NUMBER: US/09/839,479
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: US 09/418,710
;  PRIOR FILING DATE: 1999-10-15
;  PRIOR APPLICATION NUMBER: PCT/JP98/01783
;  PRIOR FILING DATE: 1998-04-17
;  PRIOR APPLICATION NUMBER: JP 9/310027
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: JP 9/116570
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
;   LENGTH: 1531
;   TYPE: PRT
```

```
;   ORGANISM: Homo sapiens
US-09-839-479-29

  Query Match             5.1%;  Score 147;  DB 4;  Length 1531;
  Best Local Similarity   36.7%;  Pred. No. 0.00022;
  Matches   29;  Conservative   13;  Mismatches   29;  Indels    8;  Gaps    2;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : : || ::|||: |||::|| | |  || : || || |:        |  |
Db       1191 CKVCPKKGEDDKLILCDECNKAFHLFCLRPALYEVP-DGEWQCPACQP-------ATARR 1242

Qy        116 RESKKNAKMASATSSSQRD 134
                :|    ||:  |: |
Db       1243 NSRGRNYTEESASEDSEDD 1261


RESULT 13
US-09-949-016-7037
; Sequence 7037, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7037
;   LENGTH: 1540
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-7037

  Query Match             5.1%;  Score 147;  DB 4;  Length 1540;
  Best Local Similarity   36.7%;  Pred. No. 0.00023;
  Matches   29;  Conservative   13;  Mismatches   29;  Indels    8;  Gaps    2;

Qy         56 CHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERL 115
              | :|  : : || ::||||: |||::|| | |  || : || || |:        |  |
Db       1244 CKVCRKKGEDDKLILCDECNKAFHLFCLRPALYEVP-DGEWQCPACQP-------ATARR 1295

Qy        116 RESKKNAKMASATSSSQRD 134
                :|    ||:  |: |
Db       1296 NSRGRNYTEESASEDSEDD 1314


RESULT 14
US-09-538-092-1033
; Sequence 1033, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
;  APPLICANT: Giot, Loic
;  APPLICANT:  Mansfield, Traci A.
;  TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;  FILE REFERENCE: 15966-542
;  CURRENT APPLICATION NUMBER: US/09/538,092
;  CURRENT FILING DATE:  2000-03-29
;  PRIOR APPLICATION NUMBER: 60/127,352
;  PRIOR FILING DATE: 1999-04-01
;  PRIOR APPLICATION NUMBER: 60/178,965
;  PRIOR FILING DATE: 2000-02-01
;  NUMBER OF SEQ ID NOS: 1387
;  SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1033
;   LENGTH: 1722
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (0)...(0)
;   OTHER INFORMATION: Polypeptide Accession Number P29375
US-09-538-092-1033

  Query Match             5.1%;  Score 147;  DB 4;  Length 1722;
  Best Local Similarity   34.8%;  Pred. No. 0.00027;
  Matches   31;  Conservative   13;  Mismatches   37;  Indels    8;  Gaps    2;

Qy         13 DEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCGGRQDPDKQLMCD 72
              ||| :  :    |   :  ||:: |    :  |       :  |  ||   : || |:||
Db       260 DEVTRRRKVTNRSDAFNMQMRQRKGTLSVNFVD-------LYVCMFCGRGNNEDKLLLCD 312

Qy         73 ECDMAFHIYCLDPPLSSVPSEDEWYCPEC 101
              || ::| :|| |||  ||  | | ||:|
Db       313 GCDDSYHTFCLIPPLPDVPKGD-WRCPKC 340


RESULT 15
US-09-051-019-2
; Sequence 2, Application US/09051019
; Patent No. 6103229
;  GENERAL INFORMATION:
;    APPLICANT:  KAHMANN, Regine and QUADBECK-SEEGER, Claudia
;    TITLE OF INVENTION:  Regulatory gene from Ustilago maydis
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Keil & Weinkauf
;      STREET:  1101 Connecticut Avenue
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.2 Mb storage
;      COMPUTER:  IBM AT-compatible, Pentium processor
;      OPERATING SYSTEM:  Windows 98
;      SOFTWARE:  WordPerfect version 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/051,019
;      FILING DATE:  31-MAR-1998
;      CLASSIFICATION:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2289 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-09-051-019-2

  Query Match             5.1%;  Score 146.5;  DB 3;  Length 2289;
  Best Local Similarity   34.5%;  Pred. No. 0.00046;
  Matches   29;  Conservative   15;  Mismatches   31;  Indels    9;  Gaps    3;

Qy         25 SPMVDNPMRRKSGPSCKHCKDDV-------NRLCRVCACHLCGGRQDPDKQLMCDECDMA 77
              :||     ||: | |  | : |        |:      | :|   :|    |:||||:
Db       505 TPMTPLISRRRKGVS-PHLEADSYLRAQAGNQAQEEQMCEICLRGEDGPNMLLCDECNRG 563

Qy         78 FHIYCLDPPLSSVPSEDEWYCPEC 101
              :|:||| | |:|:| : :|:|| |
Db       564 YHMYCLQPALTSIP-KSQWFCPPC 586


Search completed: July 25, 2005, 17:08:21
Job time : 45 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:38:02 ; Search time 166 Seconds
(without alignments)
1237.167 Million cell updates/sec

Title: US-10-019-071-2_COPY_263_793
Perfect score: 2896
Sequence: 1 DDSLNDCRIIFVDEVFKIER.....VNQPLQTVLNQLFPGYGNGR 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2896	100.0	793	4	ABB76980	Abb76980 Human Inv
2	2887	99.7	793	6	ABR48157	Abr48157 Human bla
3	2887	99.7	793	7	ADF61820	Adf61820 Human NP9
4	2887	99.7	793	7	ADF76781	Adf76781 Novel hum
5	2887	99.7	793	8	ADN05229	Adn05229 Antipsori
6	2887	99.7	793	8	ADO20357	Ado20357 Human PRO
7	2887	99.7	793	8	ADP55676	Adp55676 Human PRO
8	2815	97.2	780	6	ABU56628	Abu56628 Lung canc
9	2815	97.2	780	7	ADN38838	Adn38838 Cancer/an
10	1718.5	59.3	802	6	ABR82238	Abr82238 Human nuc
11	1562	53.9	770	8	ABM83220	Abm83220 Human dia
12	1297	44.8	645	6	ABU69599	Abu69599 Human NF-
13	1297	44.8	645	6	ABU69620	Abu69620 Human NF-
14	926	32.0	174	4	ABB76983	Abb76983 Human Inv
15	823.5	28.4	198	5	ABP64013	Abp64013 Human ORF
16	672	23.2	133	4	AAU16348	Aau16348 Human nov
17	672	23.2	133	6	ABU55417	Abu55417 Human nov
18	588.5	20.3	617	7	ADD30490	Add30490 Plant yie
19	588.5	20.3	617	8	ADI43851	Adi43851 Plant tra
20	428.5	14.8	233	3	AAB42314	Aab42314 Human ORF
21	420.5	14.5	178	4	AAU16350	Aau16350 Human nov
22	420.5	14.5	178	6	ABU55419	Abu55419 Human nov
23	420.5	14.5	180	4	AAM42002	Aam42002 Human pol
24	357.5	12.3	110	4	AAU15895	Aau15895 Human nov
25	357.5	12.3	110	6	ABU54964	Abu54964 Human nov
26	287	9.9	71	4	AAM40216	Aam40216 Human pol
27	287	9.9	71	4	AAB73737	Aab73737 Ring fing
28	263	9.1	781	3	AAG32044	Aag32044 Arabidops
29	263	9.1	785	3	AAG32043	Aag32043 Arabidops
30	263	9.1	788	3	AAG32042	Aag32042 Arabidops
31	247	8.5	89	4	AAO07182	Aao07182 Human pol
32	235.5	8.1	670	8	ADN72155	Adn72155 Thale cre
33	235.5	8.1	1072	3	AAG41664	Aag41664 Arabidops
34	235.5	8.1	1079	3	AAG41663	Aag41663 Arabidops
35	235.5	8.1	1132	3	AAG41662	Aag41662 Arabidops
36	229.5	7.9	856	8	ADM48182	Adm48182 Polypepti
37	190	6.6	87	5	ABB97911	Abb97911 Human sec
38	186.5	6.4	1398	5	ABG97491	Abg97491 Human NOV
39	186.5	6.4	1400	4	AAB83348	Aab83348 AAP-2 pro
40	186.5	6.4	1445	6	ABR41365	Abr41365 Human DIT
41	180.5	6.2	2897	4	ABB58514	Abb58514 Drosophil
42	177	6.1	4019	4	AAE13839	Aae13839 Human lun
43	177	6.1	4019	7	ADD66733	Add66733 Human lun
44	177	6.1	4019	7	ADE87987	Ade87987 Human lun
45	177	6.1	4025	5	ABP69736	Abp69736 Human pol

ALIGNMENTS

RESULT 1
ABB76980
ID ABB76980 standard; protein; 793 AA.
XX
AC ABB76980;
XX
DT 22-JUL-2002 (first entry)
XX
DE Human Inverted CCAAT box binding protein, ICBP90.
XX
KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;
KW cell proliferation control; inverted CCAAT box; cancer.
XX
OS Homo sapiens.
XX
PN WO200078949-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-FR001747.
XX
PR 22-JUN-1999; 99FR-00007935.
XX
PA (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.
XX
PI Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;
XX
DR WPI; 2001-091571/10.
DR N-PSDB; ABL58020.
XX
PT Novel inverted CCAAT box binding protein, and related nucleic acids,
PT antibodies and specific ligands, useful for treating and preventing
PT cancer.
XX
PS Claim 1; Fig 7; 115pp; French.
XX
CC The present sequence is the protein sequence for human ICBP90 (inverted
CC CCAAT box binding protein). The inverted CCAAT box is implicated in cell
CC proliferation control. Several copies of the inverted CCAAT box are
CC present in the promoter of the topoisomerase IIalpha gene, and also
CC functions as a nuclear receptor. ICBP90 and its coding sequence are
CC useful for treatment and/or prevention of cancer
XX
SQ Sequence 793 AA;

Query Match 100.0%; Score 2896; DB 4; Length 793;
Best Local Similarity 100.0%; Pred. No. 2.6e-257;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793
```

RESULT 2
ABR48157
ID ABR48157 standard; protein; 793 AA.
XX
AC ABR48157;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated protein sequence SEQ ID NO:27.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR WPI; 2003-201532/19.
DR N-PSDB; ACC50965.
XX
PT Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 10; Page 238; 307pp; English.
XX
CC The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridises to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 793 AA;

Query Match 99.7%; Score 2887; DB 6; Length 793;
Best Local Similarity 99.6%; Pred. No. 1.8e-256;
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793
```

RESULT 3
ADF61820
ID ADF61820 standard; protein; 793 AA.
XX
AC ADF61820;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human NP95 protein.
XX

```
KW   cell cycle arrest; cytostatic; antipsoriatic; antiarteriosclerotic;
KW   vasotropic; antithyroid; melanoma; breast; ovarian; lung;
KW   gastrointestinal; colon cancer; Grave's disease; psoriasis;
KW   atherosclerosis; restenosis; vasoproliferative; human; NP95.
XX
OS   Homo sapiens.
XX
PN   WO2003088910-A2.
XX
PD   30-OCT-2003.
XX
PF   15-APR-2003; 2003WO-US011867.
XX
PR   15-APR-2002; 2002US-00123568.
PR   15-APR-2002; 2002US-00123731.
PR   16-APR-2002; 2002US-0373366P.
XX
PA   (RIGE-) RIGEL PHARM INC.
XX
PI   Hitoshi Y,  Jenkins Y;
XX
DR   WPI; 2003-865396/80.
DR   N-PSDB; ADF61819.
XX
PT   Identifying a compound that modulates cell cycle arrest, for treating
PT   e.g. cancer, comprises contacting a cell comprising a target polypeptide
PT   and determining the chemical or phenotypic effect of the compound upon
PT   the cell.
XX
PS   Claim 1; SEQ ID NO 4; 176pp; English.
XX
CC   The invention relates to a novel method for identifying a compound that
CC   modulates cell cycle arrest comprising contacting a target polypeptide
CC   within a cell with a compound and determining the chemical or phenotypic
CC   effect of the compound upon the cell. The method of the invention has
CC   cytostatic, antipsoriatic, antiarteriosclerotic, vasotropic and
CC   antithyroid applications and may be useful for identifying a compound
CC   that modulates cell cycle arrest. Such compounds may subsequently be used
CC   for developing therapeutic reagents to treat melanoma, breast, ovarian,
CC   lung, gastrointestinal or colon cancer, as well as other proliferative
CC   diseases such as Grave's disease, psoriasis, atherosclerosis, restenosis
CC   and other vasoproliferative diseases. The current sequence is that of the
CC   human NP95 protein of the invention.
XX
SQ   Sequence 793 AA;

  Query Match             99.7%;  Score 2887;  DB 7;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-256;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 4
ADF76781
ID   ADF76781 standard; protein; 793 AA.
XX
AC   ADF76781;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Novel human secreted and transmembrane protein SeqID 456.
XX
KW   human; PRO; membrane bound protein; membrane bound receptor;
KW   cell proliferation; cell migration; cell differentiation;
KW   mitogenic factor; survival factor; cytotoxic factor;
KW   differentiation factor; neuropeptide; hormone; cell receptor;
KW   receptor-ligand interaction; cytostatic; chondrocyte; tumour.
XX
OS   Homo sapiens.
XX
PN   WO2003072035-A2.
XX
PD   04-SEP-2003.
XX
PF   21-FEB-2003; 2003WO-US005241.
XX
PR   22-FEB-2002; 2002US-0359461P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary SC,  Clark H,  Hunte B,  Jackman JK,  Schoenfeld JR;
PI   Williams PM,  Wood WI,  Wu TD;
XX
DR   WPI; 2003-721702/68.
DR   N-PSDB; ADF76780.
XX
PT   New PRO polypeptides, useful for diagnosing and treating an immune
PT   related disorder, e.g. systemic lupus erythematosus, rheumatoid
PT   arthritis, osteoarthritis, juvenile chronic arthritis, thyroiditis or
PT   diabetes mellitus.
XX
PS   Claim 10; SEQ ID NO 456; 918pp; English.
XX
CC   This invention relates to novel nucleic acids encoding human PRO secreted
CC   and transmembrane proteins. Extracellular proteins play important roles
CC   in the formation, differentiation and maintenance of multicellular
CC   organisms. The fate of many individual cells (for example proliferation,
CC   migration or differentiation) is typically governed by information
CC   received from other cells and the immediate environment. The information
CC   is often transmitted by secreted polypeptides (for example mitogenic
CC   factors, survival factors, cytotoxic factors, differentiation factors,
CC   neuropeptides and hormones) which are received and interpreted by diverse
CC   cell receptors or membrane bound proteins. These membrane bound proteins
CC   and receptors may be of use as pharmaceutical and diagnostic agents, such
CC   as in the blocking of receptor-ligand interactions. The current invention
CC   provides the amino acid sequences of novel human membrane bound receptors
CC   and proteins, along with the cDNA sequences encoding them. The novel
CC   proteins of the invention may have cytostatic activities through the
CC   stimulation of chondrocytes. The nucleic acids of the invention may be
CC   useful for the manufacture of a medicament for diagnosing or treating a
CC   tumour in a mammal. In addition, they may be useful for measuring or
```

```
CC   detecting the expression of a tumour associated gene. The present
CC   sequence is the amino acid sequence of a human PRO protein of the
CC   invention.
XX
SQ   Sequence 793 AA;

  Query Match             99.7%;  Score 2887;  DB 7;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-256;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 5
ADN05229
ID   ADN05229 standard; protein; 793 AA.
XX
AC   ADN05229;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #791.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN05228.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 1623; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 793 AA;

  Query Match             99.7%;  Score 2887;  DB 8;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-256;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 6
ADO20357
ID   ADO20357 standard; protein; 793 AA.
XX
AC   ADO20357;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human PRO polypeptide #626.
XX
KW   Human; PRO; immune related disorder; systemic lupus erythematosus;
KW   rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW   systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW   autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
```

KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
PN WO2004043361-A2.
XX
PD 27-MAY-2004.
XX
PF 06-NOV-2003; 2003WO-US035268.
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
PA (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-420067/39.
DR N-PSDB; ADO20356.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 7; SEQ ID NO 1252; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 793 AA;

Query Match 99.7%; Score 2887; DB 8; Length 793;
Best Local Similarity 99.6%; Pred. No. 1.8e-256;
Matches 529; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622
```

```
Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793
```

RESULT 7
ADP55676
ID ADP55676 standard; protein; 793 AA.
XX
AC ADP55676;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human PRO protein sequence SEQ ID NO:1652.
XX
KW human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004039956-A2.
XX
PD 13-MAY-2004.
XX
PF 28-OCT-2003; 2003WO-US034381.
XX
PR 29-OCT-2002; 2002US-0422472P.
XX
PA (GETH) GENENTECH INC.
XX
PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-376182/35.
DR N-PSDB; ADP55675.
XX
PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT and treating an immune related disease, e.g. systemic lupus
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT stimulating an immune response.
XX
PS Claim 1; SEQ ID NO 1652; 3009pp; English.
XX
CC The present invention describes an isolated PRO nucleic acid (I). Also
CC described: (1) a vector comprising (I); (2) a host cell comprising the
CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC antibody which specifically binds to a polypeptide of (4); (7) a
CC composition of matter comprising a polypeptide of (4), an agonist or
CC antagonist of the polypeptide or an antibody that binds to the
CC polypeptide in combination with a carrier; (8) an article of manufacture
CC comprising a container, a label on the container and a composition of
CC matter of (7); (9) a method of treating an immune related disease in a
CC mammal; (10) a method for determining the presence of a PRO polypeptide
CC in a sample suspected of having the polypeptide; (11) a method of
CC diagnosing an immune related disease or an inflammatory immune response
CC in mammal; (12) a method of identifying a compound that inhibits or
CC mimics the activity of or expression of a gene encoding a PRO polypeptide
CC ; and (13) a method of stimulating the immune response in a mammal. The

```
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO protein from the present invention.
XX
SQ   Sequence 793 AA;

  Query Match             99.7%;  Score 2887;  DB 8;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-256;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793


RESULT 8
ABU56628
ID   ABU56628 standard; protein; 780 AA.
XX
AC   ABU56628;
XX
DT   02-APR-2003  (first entry)
XX
DE   Lung cancer-associated polypeptide #221.
XX
KW   Lung cancer-associated polypeptide; cytostatic; emphysema;
KW   antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW   small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW   chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW   interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS   Unidentified.
XX
PN   WO200286443-A2.
XX
PD   31-OCT-2002.
XX
PF   18-APR-2002; 2002WO-US012476.
XX
PR   18-APR-2001; 2001US-0284770P.
PR   10-MAY-2001; 2001US-0290492P.
PR   09-NOV-2001; 2001US-0339245P.
PR   13-NOV-2001; 2001US-0350666P.
PR   29-NOV-2001; 2001US-0334370P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Aziz N,  Murray R;
XX
DR   WPI; 2003-093161/08.
DR   N-PSDB; ABX76357.
XX
PT   Detecting a lung cancer-associated transcript in a cell from a patient
PT   for treating lung cancer, by contacting a biological sample from the
PT   patient with a polynucleotide that exhibits increased or decreased
PT   expression in lung cancer.
XX
PS   Claim 27; Page 357; 453pp; English.
XX
CC   The invention relates to a method for detecting a lung cancer-associated
CC   transcript in a cell from a patient, comprising contacting a biological
CC   sample from the patient with a polynucleotide that selectively hybridises
CC   to a sequence that is at least 80 % identical to a gene that exhibits
CC   increased or decreased expression in lung cancer samples. Lung cancer-
CC   associated polynucleotides and polypeptides are used for identifying a
CC   compound that modulates a lung cancer-associated polypeptide, for
CC   inhibiting proliferation of a lung cancer-associated cell to treat lung
CC   cancer in a patient and for treating a mammal having lung cancer by
CC   administering a modulatory compound identified. The methods are useful
CC   for treating lung cancer, such as small cell lung cancer, non-small cell
CC   lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC   emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC   hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC   bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC   for diagnostic purposes and as targets for screening for therapeutic
CC   compounds that modulate lung cancer, such as antibodies. Sequences
CC   ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
CC   invention
XX
SQ   Sequence 780 AA;

  Query Match             97.2%;  Score 2815;  DB 6;  Length 780;
  Best Local Similarity   99.6%;  Pred. No. 7.6e-250;
  Matches  516;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              ||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 518
                |||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 780


RESULT 9
ADN38838
ID   ADN38838 standard; protein; 780 AA.
XX
AC   ADN38838;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:156.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO2003042661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D,  Aziz N,  Ginsburg WM,  Gish KC,  Glynne R,  Hevezi PA;
PI   Mack DH,  Murray R,  Watson SR,  Wilson KE,  Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN38837.
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO 156; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 780 AA;

  Query Match             97.2%;  Score 2815;  DB 7;  Length 780;
  Best Local Similarity   99.6%;  Pred. No. 7.6e-250;
  Matches  516;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
                |||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy        361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy        421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy        481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 518
                |||||||||||||||||||||||||||||||||||||
Db        743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQT 780


RESULT 10
ABR82238
ID   ABR82238 standard; protein; 802 AA.
XX
AC   ABR82238;
XX
DT   13-OCT-2003  (first entry)
XX
DE   Human nucleic acid-associated protein (NAAP)-Id 4706628CD1.
XX
```

```
KW   NAAP; nucleic acid-associated protein; cardiant; cytostatic; transgenic;
KW   neuroprotective; gene therapy; human; cancer.
XX
OS   Homo sapiens.
XX
PN   WO2003052048-A2.
XX
PD   26-JUN-2003.
XX
PF   02-MAY-2002; 2002WO-US014276.
XX
PR   04-MAY-2001; 2001US-0288598P.
PR   17-MAY-2001; 2001US-0291776P.
PR   18-MAY-2001; 2001US-0292172P.
PR   25-MAY-2001; 2001US-0293564P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Yue H,  Ding L,  Baughn MR,  Lal PG,  Yue H,  Hafalia AJA,  Lee EA;
PI   Ison CH,  Becha SD,  Gururajan R,  Emerling BM,  Griffin JA,  Tang YT;
PI   Lu DAM,  Yao MG,  Chawla NK,  Ramkumar J,  Gandhi AR,  Lee SY;
PI   Richardson TW,  Yang J,  Elliott VS,  Lu Y,  Thangavelu K,  He A;
PI   Azimzai Y,  Raumann BE,  Swarnakar A,  Burford N;
XX
DR   WPI; 2003-541640/51.
DR   N-PSDB; ACF35646.
XX
PT   New human nucleic acid-associated proteins polypeptide, useful for
PT   preparing a composition for diagnosing or treating e.g., cardiovascular
PT   or neurological disorders.
XX
PS   Claim 1; Page 168-169; 212pp; English.
XX
CC   The invention relates to human nucleic acid-associated proteins (NAAP)
CC   and encoding polynucleotides. The NAAP polypeptides can be expressed by
CC   standard recombinant methodology. The polypeptides are useful for
CC   preparing a composition for diagnosing or treating a disease or condition
CC   associated with decreased expression or overexpression of functional NAAP
CC   e.g., cardiovascular or neurological disorders or cancer. The present
CC   sequence represents a human NAAP polypeptide
XX
SQ   Sequence 802 AA;

  Query Match             59.3%;  Score 1718.5;  DB 6;  Length 802;
  Best Local Similarity   59.2%;  Pred. No. 9.7e-149;
  Matches  316;  Conservative   77;  Mismatches  116;  Indels   25;  Gaps    6;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||:||||||||||||||||||: |||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: ||||||||::|:|||||    | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              |||||||   | |||||||||||||| || ||| || :| ::| | :||| ||       :
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-----PSD 644

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              :| :  | : ::|  |    |          |    || :   : :     :| : || |
Db        645 KEGKKPKGQSKKQPSGTTKRP---------ISDDDCPSASKVYKASDSAEAIEAFQLTPQ 695

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  ||||  | |||:|||: | : |        || |:|::| |:|||||::|:|| |
Db        696 QQHLIREDCQNQKLWDEVLSHLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 748

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||| |||:||||||||||:|||::| |  |: |||:|:  ||||  ||
Db        749 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYIMIPNEILQTLLDLFFPGYSKGR 802


RESULT 11
ABM83220
ID   ABM83220 standard; protein; 770 AA.
XX
AC   ABM83220;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human diagnostic and therapeutic pprotein SEQ ID NO:3469.
XX
KW   gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS   Homo sapiens.
XX
PN   WO2004023973-A2.
XX
PD   25-MAR-2004.
XX
PF   12-SEP-2003; 2003WO-US028227.
XX
PR   12-SEP-2002; 2002US-0410259P.
PR   12-SEP-2002; 2002US-0410260P.
XX
PA   (INCY-) INCYTE CORP.
XX
PI   Schmidt JP,  Wright RJ,  Bruns CM,  Marjanovic MM,  Shen F;
PI   Harthshorne TA,  Suchorolski MT,  Altus CM,  Pitts SJ,  Elder LV;
PI   Mooney EM,  Delegeane AM,  Panesar IS,  Banville SC,  Reddy TP;
PI   Stevens KA,  Blanchard JL,  Panzer SR,  Wang X,  Au AP,  Gerstin EH;
PI   Peralta CH,  Anderson SB,  Rioux P,  Shen EJ,  Wu MC,  Stuve LL;
PI   Lagace RE,  Spiro PA,  Stewart EA,  Wingrove J,  Vitt UA,  Kirton ES;
PI   Xu Y,  Kwong M,  Policky JL,  Hurwitz BL,  Ma Y,  Jackson JL,  Gietzen D;
PI   Patury S,  Shi X,  Suarez CJ;
XX
DR   WPI; 2004-329368/30.
DR   N-PSDB; ACN41872.
XX
PT   New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT   in diagnosing a condition, disease or disorder associated with human
PT   molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT   in gene mapping.
XX
PS   Claim 27; Page; 190pp; English.
XX
CC   The invention relates to novel diagnostic and therapeutic polynucleotides
CC   selected from one of the 2722 sequences defined in the specification. A
CC   polynucleotide of the invention may have a use in gene therapy. The human
CC   diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC   used to diagnose a particular condition, disease or disorder associated
CC   with human molecules, e.g. cell proliferative disorders,
CC   autoimmune/inflammatory disorder, developmental disorder, endocrine
CC   disorder, neurological disorders, gastrointestinal disorders, or
CC   infections caused by virus, bacteria, fungi or parasite. The dithp
CC   molecules may also be used in genetic mapping, in identifying individuals
CC   from minute biological samples, in detecting single nucleotide
CC   polymorphisms, as molecular weight markers, and for somatic or germline
CC   gene therapy. The present sequence represents a dithp protein of the
CC   invention. Note: The sequence data for this patent is not represented in
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ   Sequence 770 AA;
```

```
  Query Match             53.9%;  Score 1562;  DB 8;  Length 770;
  Best Local Similarity   54.5%;  Pred. No. 2.4e-134;
  Matches  291;  Conservative   75;  Mismatches  112;  Indels   56;  Gaps    7;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        290 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 348

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        349 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 408

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||:||||||||||||||||||: ||||||||||:|
Db        409 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 468

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: |||||||||::|:||||    | ||
Db        469 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 528

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:|
Db        529 LTNMNRALALNCDAPLDDKIGAESRNWRAG---------------------------- 558

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              ||||||    | |||||||||||||| || ||| || :| ::| | :||| ||       :
Db        559 -VKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-----PSD 612

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              :| :  | : ::|  |    |          |    || :   : :     :| : || |
Db        613 KEGKKPKGQSKKQPSGTTKRP---------ISDDDCPSASKVYKASDSAEAIEAFQLTPQ 663

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  |||||  | |||:|||: | : |         || |:|::| |:|||||::|:|| |
Db        664 QQHLIREDCQNQKLWDEVLSHLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 716

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||| |||:||||||||||:|||::| |  |: |||:|:  ||||  ||
Db        717 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYIMIPNEILQTLLDLFFPGYSKGR 770


RESULT 12
ABU69599
ID   ABU69599 standard; protein; 645 AA.
XX
AC   ABU69599;
XX
DT   05-JUN-2003  (first entry)
XX
DE   Human NF-kappaB associated polypeptide sequence #5.
XX
KW   Human; nuclear factor-kappaB; NF-kappaB; immune disorder; cancer;
KW   inflammatory disorder; apoptosis; hepatic disorder; Hodgkin's lymphoma;
KW   haematopoietic tumour; hyper-IgM syndrome; viral infection; asthma;
KW   hypohidrotic ectodermal dysplasia; human immunodeficiency virus; HIV;
KW   X-linked anhidrotic ectodermal dysplasia; al incontinentia pigmenti;
KW   influenza; rheumatoid arthritis; inflammatory bowel disease; colitis;
KW   atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW   experimental allergic encephalomyelitis; autoimmune disorder; wound;
KW   hyper immune activity; acute phase response; hypercongenital condition;
KW   birth defect; necrotic lesion; organ transplant rejection; pancreas;
KW   signal transduction; hyperproliferative disorder; diabetes mellitus;
KW   vitamin B12 malabsorption; neurological disorder; Huntington's chorea;
KW   Turner's syndrome; bacterial infection; cardiovascular disorder;
KW   infertility; psoriasis; haemolytic anaemia; antiinflammatory; anti-HIV;
KW   cytostatic; hepatotropic; virucide; antirheumatic; antiarthritic;
KW   antiasthmatic; immunomodulator; antidiabetic; antiallergic;
KW   neuroprotective; immunosuppressive; vulnerary; antibacterial;
KW   antiinfertility; antianaemic; antipsoriatic; cerebroprotective; cardiant;
KW   antiarteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   WO200286076-A2.
XX
PD   31-OCT-2002.
XX
PF   19-APR-2002; 2002WO-US012636.
XX
PR   19-APR-2001; 2001US-0284962P.
PR   26-APR-2001; 2001US-0286645P.
PR   09-JAN-2002; 2002US-0346986P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Carman J,  Feder J,  Nadler S;
XX
DR   WPI; 2003-093119/08.
DR   N-PSDB; ACA54638.
XX
PT   Novel NF-kappaB-associated polypeptides and polynucleotides useful for
PT   diagnosing, treating and preventing cancer, hepatic disorders, aberrant
PT   apoptosis, viral infections, autoimmune disorders, asthma and stroke.
XX
PS   Claim 6; Page 493-494; 608pp; English.
XX
CC   The present invention relates to the isolation of human nuclear factor-
CC   kappaB (NF-kappaB) associated polypeptides and polynucleotides. The NF-
CC   kappaB associated polypeptide and polynucleotide sequences are useful for
CC   preventing, treating or ameliorating various disorders including immune
CC   disorders, inflammatory disorders, cancers, disorders relating to
CC   aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas,
CC   haematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal
CC   dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, al
CC   incontinentia pigmenti, viral infections (e.g. those caused by human
CC   immunodeficiency virus (HIV), human T-cell lymphotropic virus (HTLV),
CC   hepatitis B, hepatitis C, Epstein Barr virus (EBV), influenza),
CC   rheumatoid arthritis, inflammatory bowel disease, colitis, asthma,
CC   atherosclerosis, cachexia, euthyroid sick syndrome, stroke, experimental
CC   allergic encephalomyelitis (EAE), autoimmune disorders, disorders related
CC   to hyper immune activity, disorders related to aberrant acute phase
CC   responses, hypercongenital conditions, birth defects, necrotic lesions,
CC   wounds, organ transplant rejection, disorders related to aberrant signal
CC   transduction, hyperproliferative disorders, diseases of the pancreas
CC   (e.g. diabetes mellitus, vitamin B12 malabsorption), neurological
CC   disorders (e.g. Huntington's chorea), Turner's syndrome, bacterial
CC   infections, cardiovascular disorders, infertility, psoriasis and
CC   haemolytic anaemia. The present sequence represents a human NF-kappaB
CC   associated polypeptide of the invention
XX
SQ   Sequence 645 AA;

  Query Match             44.8%;  Score 1297;  DB 6;  Length 645;
  Best Local Similarity   67.2%;  Pred. No. 5.1e-110;
  Matches  231;  Conservative   48;  Mismatches   61;  Indels    4;  Gaps    3;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||:||||||||||||||||||: ||||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: |||||||||::|:||||    | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:|||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589
```

```
Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGL 341
              |||||||   | |||||||||||| || ||| || :| ::| |
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCL 633
```

RESULT 13
ABU69620
ID ABU69620 standard; protein; 645 AA.
XX
AC ABU69620;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human NF-kappaB associated polypeptide sequence #23.
XX
KW Human; nuclear factor-kappaB; NF-kappaB; immune disorder; cancer;
KW inflammatory disorder; apoptosis; hepatic disorder; Hodgkin's lymphoma;
KW haematopoietic tumour; hyper-IgM syndrome; viral infection; asthma;
KW hypohidrotic ectodermal dysplasia; human immunodeficiency virus; HIV;
KW X-linked anhidrotic ectodermal dysplasia; al incontinentia pigmenti;
KW influenza; rheumatoid arthritis; inflammatory bowel disease; colitis;
KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW experimental allergic encephalomyelitis; autoimmune disorder; wound;
KW hyper immune activity; acute phase response; hypercongenital condition;
KW birth defect; necrotic lesion; organ transplant rejection; pancreas;
KW signal transduction; hyperproliferative disorder; diabetes mellitus;
KW vitamin B12 malabsorption; neurological disorder; Huntington's chorea;
KW Turner's syndrome; bacterial infection; cardiovascular disorder;
KW infertility; psoriasis; haemolytic anaemia; antiinflammatory; anti-HIV;
KW cytostatic; hepatotropic; virucide; antirheumatic; antiarthritic;
KW antiasthmatic; immunomodulator; antidiabetic; antiallergic;
KW neuroprotective; immunosuppressive; vulnerary; antibacterial;
KW antiinfertility; antianaemic; antipsoriatic; cerebroprotective; cardiant;
KW antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO200286076-A2.
XX
PD 31-OCT-2002.
XX
PF 19-APR-2002; 2002WO-US012636.
XX
PR 19-APR-2001; 2001US-0284962P.
PR 26-APR-2001; 2001US-0286645P.
PR 09-JAN-2002; 2002US-0346986P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Carman J, Feder J, Nadler S;
XX
DR WPI; 2003-093119/08.
DR N-PSDB; ACA54638.
XX
PT Novel NF-kappaB-associated polypeptides and polynucleotides useful for
PT diagnosing, treating and preventing cancer, hepatic disorders, aberrant
PT apoptosis, viral infections, autoimmune disorders, asthma and stroke.
XX
PS Claim 4; Page 497-499; 608pp; English.
XX
CC The present invention relates to the isolation of human nuclear factor-
CC kappaB (NF-kappaB) associated polypeptides and polynucleotides. The NF-
CC kappaB associated polypeptide and polynucleotide sequences are useful for
CC preventing, treating or ameliorating various disorders including immune
CC disorders, inflammatory disorders, cancers, disorders relating to
CC aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas,
CC haematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal
CC dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, al
CC incontinentia pigmenti, viral infections (e.g. those caused by human
CC immunodeficiency virus (HIV), human T-cell lymphotropic virus (HTLV),
CC hepatitis B, hepatitis C, Epstein Barr virus (EBV), influenza),
CC rheumatoid arthritis, inflammatory bowel disease, colitis, asthma,
CC atherosclerosis, cachexia, euthyroid sick syndrome, stroke, experimental
CC allergic encephalomyelitis (EAE), autoimmune disorders, disorders related
CC to hyper immune activity, disorders related to aberrant acute phase
CC responses, hypercongenital conditions, birth defects, necrotic lesions,
CC wounds, organ transplant rejection, disorders related to aberrant signal
CC transduction, hyperproliferative disorders, diseases of the pancreas
CC (e.g. diabetes mellitus, vitamin B12 malabsorption), neurological
CC disorders (e.g. Huntington's chorea), Turner's syndrome, bacterial
CC infections, cardiovascular disorders, infertility, psoriasis and
CC haemolytic anaemia. The present sequence represents a human NF-kappaB
CC associated polypeptide of the invention
XX
SQ Sequence 645 AA;

```
  Query Match             44.8%;  Score 1297;  DB 6;  Length 645;
  Best Local Similarity   67.2%;  Pred. No. 5.1e-110;
  Matches  231;  Conservative   48;  Mismatches   61;  Indels    4;  Gaps    3;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :||||:|| |||:|||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| |||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||:|||||||||||||||||||: |||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: |||||||::|:|||| | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGL 341
              |||||||   | |||||||||||| || ||| || :| ::| |
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCL 633
```

RESULT 14
ABB76983
ID ABB76983 standard; protein; 174 AA.
XX
AC ABB76983;
XX
DT 22-JUL-2002 (first entry)
XX
DE Human Inverted CCAAT box binding protein, ICBP90, fragment #3.
XX
KW Human; inverted CCAAT box binding protein; ICBP90; cytostatic;
KW cell proliferation control; inverted CCAAT box; cancer.
XX
OS Homo sapiens.
XX
PN WO200078949-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-FR001747.
XX
PR 22-JUN-1999; 99FR-00007935.
XX
PA (ADER-) ADEREGEM ASSOC DEV RECH EN GENETIQ.
XX
PI Bronner C, Hopfner R, Mousli M, Jeltsch J, Lutz Y, Oudet P;
XX
DR WPI; 2001-091571/10.
DR N-PSDB; ABL58023.
XX

```
PT   Novel inverted CCAAT box binding protein, and related nucleic acids,
PT   antibodies and specific ligands, useful for treating and preventing
PT   cancer.
XX
PS   Claim 2; Page 103; 115pp; French.
XX
CC   The present sequence is a protein fragment of human ICBP90 (inverted
CC   CCAAT box binding protein). The inverted CCAAT box is implicated in cell
CC   proliferation control. Several copies of the inverted CCAAT box are
CC   present in the promoter of the topoisomerase IIalpha gene, and also
CC   functions as a nuclear receptor. ICBP90 (ABB76980) and its coding
CC   sequence (ABL58020) are useful for treatment and/or prevention of cancer
XX
SQ   Sequence 174 AA;

  Query Match            32.0%;  Score 926;  DB 4;  Length 174;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-76;
  Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 60

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 120

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 174


RESULT 15
ABP64013
ID   ABP64013 standard; protein; 198 AA.
XX
AC   ABP64013;
XX
DT   04-NOV-2002  (first entry)
XX
DE   Human ORF383.
XX
KW   Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW   Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW   human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW   cancer; cardiovascular disease; allergy; autoimmune disease;
KW   wound healing; blood coagulation disorder; inflammatory disorder.
XX
OS   Homo sapiens.
XX
PN   US2002082206-A1.
XX
PD   27-JUN-2002.
XX
PF   30-MAY-2001; 2001US-00867550.
XX
PR   30-MAY-2000; 2000US-0208427P.
XX
PA   (LEAC/) LEACH M D.
PA   (MEHR/) MEHRABAN F.
PA   (CONL/) CONLEY P B.
PA   (TOPP/) TOPPER J N.
PA   (LAWD/) LAW D.
XX
PI   Leach MD,  Mehraban F,  Conley PB,  Topper JN,  Law D;
XX
DR   WPI; 2002-626554/67.
DR   N-PSDB; ABQ98576.
XX
PT   New polypeptide designated ORFX are present in human atherogenic cells
PT   and are useful to prevent and treat ORFX-associated disorders including
PT   cancer, allergy, wound healing or autoimmune, cardiovascular or
PT   inflammatory disease.
XX
PS   Claim 10; SEQ ID NO 766; 78pp; English.
XX
CC   The present invention relates to novel human ORFX polypeptides and their
CC   coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC   were discovered in human atherogenic cells, in particular in platelets
CC   and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC   many other tissues as well. Atherogenic cells are cells which have the
CC   potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC   nucleic acids are useful for treating or preventing a pathological
CC   condition associated with an ORFX-associated disorder, e.g. cancer,
CC   cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC   coagulation disorders or inflammatory disorders. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from the USPTO web site at
CC   seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ   Sequence 198 AA;

  Query Match            28.4%;  Score 823.5;  DB 5;  Length 198;
  Best Local Similarity  75.8%;  Pred. No. 4e-67;
  Matches  150;  Conservative   22;  Mismatches   25;  Indels    1;  Gaps    1;

Qy        122 AKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVHR 181
              ||| ||:: |:||||:||||||||:||||||||||||||||||: |||||||||  |||
Db          1 AKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEEXVHR 60

Qy        182 PHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTN 241
              ||| |||||||||:||||||||: |:|| |: |||||||::|:|:||||    | || |||
Db         61 PHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQTLTN 120

Qy        242 TNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVK 301
              |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||||||
Db        121 MNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYKVVK 180

Qy        302 YWPEKGKS-GFLVWRYLL 318
              ||||   | |||||||||
Db        181 YWPEISSSHGFLVWRYLL 198


Search completed: July 25, 2005, 17:03:40
Job time : 170 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 25, 2005, 16:39:26 ; Search time 179 Seconds
(without alignments)
1519.073 Million cell updates/sec

Title: US-10-019-071-2_COPY_263_793
Perfect score: 2896
Sequence: 1 DDSLNDCRIIFVDEVFKIER.....VNQPLQTVLNQLFPGYGNGR 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2896	100.0	793	2	Q9P115	Q9p115 homo sapien
2	2887	99.7	793	2	Q96T88	Q96t88 homo sapien
3	2212	76.4	829	2	Q7TPK1	Q7tpk1 rattus norv
4	2169	74.9	782	2	Q8VDF2	Q8vdf2 mus musculu
5	2165	74.8	782	2	Q9Z1H6	Q9z1h6 mus musculu
6	2069.5	71.5	775	2	Q6PEI0	Q6pei0 brachydanio
7	2062	71.2	776	2	Q6DRP6	Q6drp6 brachydanio
8	1758.5	60.7	803	2	Q7TMI3	Q7tmi3 mus musculu
9	1757.5	60.7	803	2	Q8K1I5	Q8k1i5 mus musculu
10	1718.5	59.3	802	2	Q96PU4	Q96pu4 homo sapien
11	1544	53.3	597	2	Q6IP39	Q6ip39 xenopus lae
12	1207	41.7	299	2	Q8C6F1	Q8c6f1 mus musculu
13	1005	34.7	189	2	Q9H6S6	Q9h6s6 homo sapien
14	886	30.6	474	2	Q8VIA1	Q8via1 mus musculu
15	816.5	28.2	516	2	Q8BJP6	Q8bjp6 mus musculu
16	792.5	27.4	503	2	Q8TAG7	Q8tag7 homo sapien
17	732	25.3	139	2	Q8J022	Q8j022 homo sapien
18	608	21.0	645	2	Q8VYZ0	Q8vyz0 arabidopsis
19	600	20.7	641	2	Q9FVS3	Q9fvs3 arabidopsis
20	591.5	20.4	617	2	Q67XP0	Q67xp0 arabidopsis
21	588.5	20.3	617	2	Q9FKA7	Q9fka7 arabidopsis
22	560	19.3	615	2	Q6NQ90	Q6nq90 arabidopsis
23	556	19.2	622	2	Q9C8E1	Q9c8e1 arabidopsis
24	556	19.2	623	2	Q680I0	Q680i0 arabidopsis
25	553.5	19.1	765	2	Q7XW58	Q7xw58 oryza sativ
26	533.5	18.4	598	2	Q9C8E0	Q9c8e0 arabidopsis
27	524	18.1	789	2	Q75M36	Q75m36 oryza sativ
28	514.5	17.8	610	2	Q9FW25	Q9fw25 oryza sativ
29	473	16.3	650	2	Q9FVS2	Q9fvs2 arabidopsis
30	427	14.7	461	2	O81463	O81463 arabidopsis
31	424	14.6	465	2	Q681I0	Q681i0 arabidopsis
32	420.5	14.5	167	2	Q659C8	Q659c8 homo sapien
33	353.5	12.2	432	2	O22280	O22280 arabidopsis
34	335	11.6	434	2	Q8H9A3	Q8h9a3 arabis gemm
35	311.5	10.8	299	2	Q9RU61	Q9ru61 deinococcus
36	310.5	10.7	301	2	Q9ADD4	Q9add4 streptomyce
37	279	9.6	684	2	Q7XHM7	Q7xhm7 oryza sativ
38	270.5	9.3	794	1	SUV5_ARATH	O82175 arabidopsis
39	263	9.1	48	2	Q9P1U7	Q9p1u7 homo sapien
40	263	9.1	790	1	SUV6_ARATH	Q8vz17 arabidopsis
41	252.5	8.7	624	1	SUV4_ARATH	Q8gzb6 arabidopsis
42	248.5	8.6	766	2	Q8H6E0	Q8h6b0 zea mays (m
43	247.5	8.5	921	2	Q6K4E6	Q6k4e6 oryza sativ
44	237.5	8.2	908	2	Q6ZJS9	Q6zjs9 oryza sativ
45	235.5	8.1	670	1	SUV1_ARATH	Q9ff80 arabidopsis

ALIGNMENTS

RESULT 1
Q9P115
ID Q9P115 PRELIMINARY; PRT; 793 AA.
AC Q9P115;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription factor ICBP90.
GN Name=ICBP90;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20110783; PubMed=10646863;
RA Hopfner R., Mousli M., Jeltsch J.M., Voulgaris A., Lutz Y., Marin C.,
RA Bellocq J.P., Oudet P., Bronner C.;
RT "ICBP90, a novel human CCAAT binding protein, involved in the
RT regulation of topoisomerase IIa expression.";
RL Cancer Res. 60:121-128(2000).
DR EMBL; AF129507; AAF28469.1; -.
DR HSSP; Q9UIG0; 1F62.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0003700; F:transcription factor activity; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; TAS.
DR InterPro; IPR011011; FYVE_PHD_ZnF.
DR InterPro; IPR003105; SRA_YDG.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 2.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS50089; ZF_RING_2; 2.
SQ SEQUENCE 793 AA; 89815 MW; D9B4161E892BB014 CRC64;

Query Match 100.0%; Score 2896; DB 2; Length 793;
Best Local Similarity 100.0%; Pred. No. 5.5e-191;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322
```

```
Qy          61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy         121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         383 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy         181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         443 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy         241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy         301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy         361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy         421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy         481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793
```

RESULT 2
Q96T88
ID Q96T88 PRELIMINARY; PRT; 793 AA.
AC Q96T88;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nuclear zinc finger protein Np95.
GN Name=UHRF1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF274048; AAK55744.1; -.
DR HSSP; Q9UIG0; 1F62.
DR Genew; HGNC:12556; UHRF1.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011011; FYVE_PHD_ZnF.
DR InterPro; IPR003105; SRA_YDG.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 2.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS50089; ZF_RING_2; 2.
SQ SEQUENCE 793 AA; 89813 MW; E65B15657525C89F CRC64;

```
  Query Match             99.7%;  Score 2887;  DB 2;  Length 793;
  Best Local Similarity   99.6%;  Pred. No. 2.3e-190;
  Matches  529;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         263 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 322

Qy          61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         323 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 382

Qy         121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         383 KAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 442

Qy         181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
               ||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db         443 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 502

Qy         241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         503 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 562

Qy         301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         563 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 622

Qy         361 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         623 ENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQSS 682

Qy         421 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         683 LIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQHN 742

Qy         481 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         743 VCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 793
```

RESULT 3
Q7TPK1
ID Q7TPK1 PRELIMINARY; PRT; 829 AA.
AC Q7TPK1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ac2-121.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Wang G.P., Chai L.Q., Yuan J.Y., Yang K.J.,
RA Yan H.M., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F., Han H.P.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY321334; AAP86266.1; -.
DR HSSP; Q9UIG0; 1F62.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.

```
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR002345; Lipocalin.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   829 AA;  93222 MW;  E6E8327F33FE74BE CRC64;

  Query Match             76.4%;  Score 2212;  DB 2;  Length 829;
  Best Local Similarity   75.9%;  Pred. No. 7.6e-144;
  Matches  404;  Conservative   58;  Mismatches   52;  Indels   18;  Gaps    7;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCG 60
              |  ||:|||||||| ||| | | ||:: :| ||||||||::|||| |: || ||||:||
Db        315 DSQLNNCRIIFVDEVLKIELPNERSPLIGSPSRRKSGPSCQYCKDDENKPCRKCACHICG 374

Qy         61 GRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKK 120
              ||: |:|||::|||||||||:||| ||| || | ||||| || |:|||| |||:|::|||
Db        375 GREAPEKQVLCDECDMAFHLYCLQPPLTCVPPEPEWYCPSCRTDSSEVVQAGEKLKKSKK 434

Qy        121 NAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVH 180
               |||||||||:|||||||||||||| |||||:||:|||||:||||||||||||||||||
Db        435 KAKMASATSSSRRDWGKGMACVGRTTECTIVPANHFGPIPGVPVGTMWRFRVQVSESGVH 494

Qy        181 RPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLT 240
              |||||||||||||:|||||||||||||:|||||||||||||||||||||| || |||||
Db        495 RPHVAGIHGRSNDGAYSLVLAGGYEDDVDNGNFFTYTGSGGRDLSGNKRTAGQSSDQKLT 554

Qy        241 NTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVV 300
              | |||||||| :||| ::|||:||| |||||||||:||||:||||||||||||||||||
Db        555 NNNRALALNCHSPIN-EKGAEAEDWRQGKPVRVVRNMKGGKHSKYAPAEGNRYDGIYKVV 613

Qy        301 KYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANREREK 360
              ||||||||||:|||||||||| || |||:|||| ::|||||||||||||||:|:
Db        614 KYWPEKGKSGFIVWRYLLRRDDTEPEPWTREGKDRTRQLGLTMQYPEGYLEALANKEK-- 671

Qy        361 ENSKREEEEQQEGGFASPRTGK-GKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQQS 419
                | ||  :  ::|    | : | || |||| |   :   || | |||:|:|||:|  ||:
Db        672 -NRKRPAKALEQG----PSSSKIGKSKRKSTGPATT---SP-RVSKKSKLEPYTLPLQQA 722

Qy        420 SLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVCQH 479
              :||:||| ||||::||:||:|      |:|:|||||:| ||||||||||||:||||||
Db        723 NLIKEDKGNAKLWDDVLSSLQD-----GPYQIFLSKVKEAFQCICCQELVFRPVTTVCQH 777

Qy        480 NVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||||||||||||||||||||   :|||||||:|||||||||:||
Db        778 NVCKDCLDRSFRAQVFSCPACRYDLDHSSPTRVNQPLQTILNQLFPGYGSGR 829


RESULT 4
Q8VDF2
ID   Q8VDF2      PRELIMINARY;      PRT;   782 AA.
AC   Q8VDF2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ubiquitin-like, containing PHD and RING finger domains, 1.
GN   Name=Uhrf1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CZECH II; TISSUE=Mammary tumor;
RA   Strausberg R.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC022167; AAH22167.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   MGD; MGI:1338889; Uhrf1.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 2.
SQ   SEQUENCE   782 AA;  88319 MW;  DC5EEDFCDF779074 CRC64;

  Query Match             74.9%;  Score 2169;  DB 2;  Length 782;
  Best Local Similarity   74.0%;  Pred. No. 6.5e-141;
  Matches  399;  Conservative   58;  Mismatches   58;  Indels   24;  Gaps    7;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMR--------RKSGPSCKHCKDDVNRLCR 52
              |  ||:|||:|||||  || | |  |:: :| :         ||||||: |||| |: ||
Db        260 DSQLNNCRIMFVDEVLMIELPKERRPLIASPSQPPPALRNTGKSGPSCRFCKDDENKPCR 319

Qy         53 VCACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG 112
               ||||:||||: |:|||:|||||||||:||| |||:||| | ||||| || |:|||| ||
Db        320 KCACHVCGGREAPEKQLLCDECDMAFHLYCLKPPLTSVPPEPEWYCPSCRTDSSEVVQAG 379

Qy        113 ERLRESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRV 172
              |:|:|||| |||||||||:|||||||||||| ||||||:||:|||||:||||||||||
Db        380 EKLKESKKKAKMASATSSSRRDWGKGMACVGRTTECTIVPANHFGPIPGVPVGTMWRFRV 439

Qy        173 QVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAE 232
              ||||||||||||||||||||:|||||||||||||:||:|||||||||||||||||||||
Db        440 QVSESGVHRPHVAGIHGRSNDGAYSLVLAGGYEDDVDNGNYFTYTGSGGRDLSGNKRTAG 499
```

```
Qy        233 QSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNR 292
              || |||||| |||||||| :||| ::||||:||| |||||||||:||||:|||||||||||
Db        500 QSSDQKLTNNNRALALNCHSPIN-EKGAEAEDWRQGKPVRVVRNMKGGKHSKYAPAEGNR 558

Qy        293 YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEA 352
              |||||||||||||:||||||||||||||| || |||:||||| ::|||||||||||||||
Db        559 YDGIYKVVKYWPERGKSGFLVWRYLLRRDDTEPEPWTREGKDRTRQLGLTMQYPEGYLEA 618

Qy        353 LANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPY 412
              |||    || |::   :  | | :| :||| | |      ||:   || | |||:|:|||
Db        619 LAN----KEKSRKRPAKALEQGPSSSKTGKSKQK----STGPT-LSSP-RASKKSKLEPY 668

Qy        413 SLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRP 472
              :|: ||::||:||| |||||::|| ||:|      |:|:|||||:| ||||||||||||
Db        669 TLSEQQANLIKEDKGNAKLWDDVLTSLQD-----GPYQIFLSKVKEAFQCICCQELVFRP 723

Qy        473 ITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              :|||||||||||||||||||||||||||:|  |   :|||||||:|||||||||:||
Db        724 VTTVCQHNVCKDCLDRSFRAQVFSCPACRYELDHSSPTRVNQPLQTILNQLFPGYGSGR 782
```

RESULT 5
Q9Z1H6
ID Q9Z1H6 PRELIMINARY; PRT; 782 AA.
AC Q9Z1H6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Nuclear protein np95 (Nuclear zinc finger protein Np95).
GN Name=Uhrf1; Synonyms=Np95;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pre-Tcell;
RX MEDLINE=99099250; PubMed=9880673;
RA Fujimori A., Matsuda Y., Takemoto Y., Hashimoto Y., Kubo E., Araki R.,
RA Fukumura R., Mita K., Tatsumi K., Muto M.;
RT "Cloning and mapping of Np95 gene which encodes a novel nuclear
RT protein associated with cell proliferation.";
RL Mamm. Genome 9:1032-1035(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; D87908; BAA74579.1; -.
DR EMBL; AF274046; AAK55743.1; -.
DR HSSP; Q9UIG0; 1F62.
DR MGD; MGI:1338889; Uhrf1.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0008283; P:cell proliferation; TAS.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR011011; FYVE_PHD_ZnF.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR003105; SRA_YDG.
DR InterPro; IPR000626; Ubiquitin.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR Pfam; PF02182; YDG_SRA; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 2.
DR SMART; SM00466; SRA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
DR PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR PROSITE; PS50016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS50089; ZF_RING_2; 2.
KW Nuclear protein.
SQ SEQUENCE 782 AA; 88303 MW; DC5EEDFCDF69619B CRC64;

```
  Query Match             74.8%;  Score 2165;  DB 2;  Length 782;
  Best Local Similarity   73.8%;  Pred. No. 1.2e-140;
  Matches  398;  Conservative   59;  Mismatches   58;  Indels   24;  Gaps    7;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMR--------RKSGPSCKHCKDDVNRLCR 52
              |  ||:|||:|||||  || | |  |:: :| :         ||||||: |||| |: ||
Db        260 DSQLNNCRIMFVDEVLMIELPKERRPLIASPSQPPPALRNTGKSGPSCRFCKDDENKPCR 319

Qy         53 VCACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG 112
               ||||:||||: |:|||:|||||||||:||| |||:||| | ||||| || |:|||| ||
Db        320 KCACHVCGGREAPEKQLLCDECDMAFHLYCLKPPLTSVPPEPEWYCPSCRTDSSEVVQAG 379

Qy        113 ERLRESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRV 172
              |:|:|||| |||||||||:|||||||||||| ||||||:||:|||||:||||||||||
Db        380 EKLKESKKKAKMASATSSSRRDWGKGMACVGRTTECTIVPANHFGPIPGVPVGTMWRFRV 439

Qy        173 QVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAE 232
              |||||||||||||||||||||:|||||||||||||:||:|||||||||||||||||||
Db        440 QVSESGVHRPHVAGIHGRSNDGAYSLVLAGGYEDDVDNGNYFTYTGSGGRDLSGNKRTAG 499

Qy        233 QSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNR 292
              || |||||| |||||||| :||| ::||||:||| |||||||||:||||:|||||||||||
Db        500 QSSDQKLTNNNRALALNCHSPIN-EKGAEAEDWRQGKPVRVVRNMKGGKHSKYAPAEGNR 558

Qy        293 YDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEA 352
              |||||||||||||:||||||||||||||| || |||:||||| ::|||||||||||||||
Db        559 YDGIYKVVKYWPERGKSGFLVWRYLLRRDDTEPEPWTREGKDRTRQLGLTMQYPEGYLEA 618

Qy        353 LANREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPY 412
              |||    || |::   :  | | :| :||| | |      ||:   || | |||:|:|||
Db        619 LAN----KEKSRKRPAKALEQGPSSSKTGKSKQK----STGPT-LSSP-RASKKSKLEPY 668

Qy        413 SLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRP 472
              :|: ||::||:||| |||||::|| ||:|      |:|:|||||:| ||||||||||||
Db        669 TLSEQQANLIKEDKGNAKLWDDVLTSLQD-----GPYQIFLSKVKEAFQCICCQELVFRP 723

Qy        473 ITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              :|||||||||||||||||||||||||||::|  |   :|||||||:|||||||||:||
Db        724 VTTVCQHNVCKDCLDRSFRAQVFSCPACRFELDHSSPTRVNQPLQTILNQLFPGYGSGR 782
```

RESULT 6
Q6PEI0
ID Q6PEI0 PRELIMINARY; PRT; 775 AA.
AC Q6PEI0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ubiquitin-like, containing PHD and RING finger domains, 1.
GN Name=uhrf1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; BC058055; AAH58055.1; -.
DR   HSSP; Q862M4; 1AAR.
DR   ZFIN; ZDB-GENE-040426-2039; uhrf1.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   775 AA;  87090 MW;  96D60695EDA2468D CRC64;

  Query Match           71.5%;  Score 2069.5;  DB 2;  Length 775;
  Best Local Similarity   69.7%;  Pred. No. 4.6e-134;
  Matches  372;  Conservative   76;  Mismatches   65;  Indels   21;  Gaps    7;

Qy          2 DSLNDCRIIFVDEVFKIERPG--EG-SPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              |||||||:|| |::||| ||  ||     |:|::| :|| || ||||  : |||| ||:
Db        258 DSLNDCRIMFVTEIYKIEEPGSAEGPGASSDSPLKRSNGPECKVCKDDPKKNCRVCNCHV 317

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              || :|||||||:|||||||| |||:|||:::| :::||||:|||||||||||||:|:||
Db        318 CGIKQDPDKQLLCDECDMAFHTYCLNPPLTTIPDDEDWYCPDCRNDASEVVLAGEKLKES 377

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || |||||:|||||||||||||||:||||||||||:||:||||:|:|||||||||
Db        378 KKKAKMASASSSSQRDWGKGMACVGRTKQCTIVPSNHYGPVPGVPVGTLWKFRVQVSESG 437

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||||||||||||:|||||||||||||| || ||||||||||||||||||||||||||
Db        438 VHRPHVAGIHGRSNDGAYSLVLAGGYEDDVDDGNEFTYTGSGGRDLSGNKRTAEQSCDQK 497

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| | :||:||||||||::||||||||: || |:|||:| :|||||||||
Db        498 LTNMNRALALNCNAAVNDKEGAEAKDWKAGKPVRVVRSSKGRKHSKYSPEDGNRYDGIYK 557

Qy        299 VVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRER 358
              |||||||||||||||||||:|:|:|  |||::||:||||||||||||||||:| :|:

Db        558 VVKYWPEKGKSGFLVWRYLLKRNDEESAPWTRDGKERIKKLGLTMQYPEGYLEAVAAKEK 617

Qy        359 EKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRR-TSKKTKVEPYSLTAQ 417
              ||||   :: |:      :|   ||| ||||          || : | || ||| | |: :
Db        618 EKENKNEDDIEE------TPT--KGKRKRKSQ--SMEEKSSPTKGTPKKMKVEAYKLSKE 667

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              | :||::|: | |||:| : ||   |        |::|||| | |||||||:|::|||| |
Db        668 QKALIKDDELNKKLWDEAMESLSLGPR-------FVNKVEEVFLCICCQEVVYQPITTEC 720

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||::|| |||:|:|::|||||:|||::| : ||:||| :| ||||||  :||
Db        721 QHNVCRECLQRSFKAEVYTCPACRHDLGKNYQVAVNKPLQAILTQLFPGYSSGR 774


RESULT 7
Q6DRP6
ID   Q6DRP6      PRELIMINARY;      PRT;   776 AA.
AC   Q6DRP6;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   NP95.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15256591; DOI=10.1073/pnas.0403929101;
RA   Amsterdam A., Nissen R.M., Sun Z., Swindell E.C., Farrington S.,
RA   Hopkins N.;
RT   "Identification of 315 genes essential for early zebrafish
RT   development.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:12792-12797(2004).
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; AY648713; AAT68031.1; -.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   776 AA;  87170 MW;  87FE9FC7A1F8664C CRC64;

  Query Match           71.2%;  Score 2062;  DB 2;  Length 776;
  Best Local Similarity   69.7%;  Pred. No. 1.5e-133;
  Matches  373;  Conservative   75;  Mismatches   65;  Indels   22;  Gaps    8;

Qy          2 DSLNDCRIIFVDEVFKIERPG--EG-SPMVDNPMRRKS-GPSCKHCKDDVNRLCRVCACH 57
              ||||||||:|| |::||| ||  ||     |:|:: :| || || ||||  : |||| ||
Db        258 DSLNDCRIMFVTEIYKIEEPGSAEGPGASSDSPLKSQSNGPECKVCKDDPKKNCRVCNCH 317
```

```
Qy         58 LCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRE 117
              :|| :|||||||:|||||||| |||:|||:::| :::||||:|||||||||||||:|:|
Db        318 VCGIKQDPDKQLLCDECDMAFHTYCLNPPLTTIPDDEDWYCPDCRNDASEVVLAGEKLKE 377

Qy        118 SKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSES 177
              ||| |||||:|||||||||||||||||:||||||||||:||:||||:|:||||||||
Db        378 SKKKAKMASASSSSQRDWGKGMACVGRTKQCTIVPSNHYGPVPGVPVGTLWKFRVQVSES 437

Qy        178 GVHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQ 237
              |||||||||||||||:|||||||||||| || ||||||||||||||||||||||||
Db        438 GVHRPHVAGIHGRSNDGAYSLVLAGGYEDDVDDGNEFTYTGSGGRDLSGNKRTAEQSCDQ 497

Qy        238 KLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIY 297
              |||| |||||||| | :||:|||||||::|||||||: || |:|||:| :|||||||
Db        498 KLTNMNRALALNCNAAVNDKEGAEAKDWKAGKPVRVVRSYKGRKHSKYSPEDGNRYDGIY 557

Qy        298 KVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              ||||||||||||||||||:|:|:|  |||::||:|||||||||||||||||:| :|
Db        558 KVVKYWPEKGKSGFLVWRYLLKRNDEESAPWTRDGKERIKKLGLTMQYPEGYLEAVAAKE 617

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRR-TSKKTKVEPYSLTA 416
              :|||| :: |:  :|  ||| |||  || : | || ||| | |:
Db        618 KEKENKNEDDIEE------TPT--KGKRKRKSQ--SMEEKSSPTKGTPKKMKVEAYKLSK 667

Qy        417 QQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTV 476
              :| :||::|: | |||:| : ||  |       |::|||| | ||||||:|::|||
Db        668 EQKALIKDDELNKKLWDEAMESLSLGPR-------FVNKVEEVFLCICCQEVVYQPITTE 720

Qy        477 CQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              ||||||::|| |||:|:|::|||||:|||::| | ||:||| :| |||||| :||
Db        721 CQHNVCRECLQRSFKAEVYTCPACRHDLGKNYQMAVNKPLQAILTQLFPGYSSGR 775
```

```
RESULT 8
Q7TMI3
ID   Q7TMI3      PRELIMINARY;      PRT;   803 AA.
AC   Q7TMI3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Np95-like ring finger protein.
GN   Name=NIRF; Synonyms=Uhrf2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Swiss Webster/NIH; TISSUE=Normal whole embryo;
RX   PubMed=14741369; DOI=10.1016/S0014-5793(03)01495-9;
RA   Mori T., Li Y., Hata H., Kochi H.;
RT   "NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a
RT   PEST-containing nuclear protein.";
RL   FEBS Lett. 557:209-214(2004).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Strausberg R.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; AB116653; BAC81739.1; -.
DR   EMBL; BC060241; AAH60241.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   803 AA;  90105 MW;  D799B0205E0E036E CRC64;

  Query Match             60.7%;  Score 1758.5;  DB 2;  Length 803;
  Best Local Similarity   60.5%;  Pred. No. 1.3e-112;
  Matches  323;  Conservative   75;  Mismatches  111;  Indels   25;  Gaps    6;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :|||||:: |||:||||:||    |:   |    ||: | | |  | :: | :|:||
Db        292 EGTLNDCRVMSVDEIFKIEKPG-AHPISFADGKFLRKNDPECDLCGGDPDKTCHMCSCHK 350

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              || ::||: ||:||||:||:||||| |||  || |: |||| |: |:|||| |||||: |
Db        351 CGEKRDPNMQLLCDECNMAYHIYCLSPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKLS 410

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:|||||||||||||||||||||||||||: |||||||||:|
Db        411 KKKAKMPSASTESRRDWGRGMACVGRTKECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 470

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||:||:|| |: |||||||::|:||||    | ||
Db        471 VHRPHVGGIHGRSNDGAYSLVLAGGFEDEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 530

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        531 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 590

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              |||||||   | ||||||||||||| || ||| || :| ::| | :||| ||
Db        591 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-------P 643

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              |||  | : : :::|  |: |      |   | :   : :     || : || |
Db        644 SEKEGKKTKGQSKKQGSEATKRPA-------SDDECPGDSKVLKASDSTDAVEAFQLTPQ 696

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  ||||  | |||:|||||| : |       || |:|::| |:||||||::|:|| |
Db        697 QQRLIREDCQNQKLWDEVLASLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 749
```

```
Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              ||.|||||| |||:||||||||||:|||::| | :|: |||:|:  ||||  ||
Db        750 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYVMVLNETLQTLLDLFFPGYSKGR 803


RESULT 9
Q8K1I5
ID   Q8K1I5      PRELIMINARY;      PRT;   803 AA.
AC   Q8K1I5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Nuclear zinc finger protein Np97.
GN   Name=Uhrf2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukemic cell line;
RA   Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; AF274047; AAM33798.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   MGD; MGI:1923718; Uhrf2.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   803 AA;  90091 MW;  99C82799305F3FA0 CRC64;

  Query Match             60.7%;  Score 1757.5;  DB 2;  Length 803;
  Best Local Similarity   60.5%;  Pred. No. 1.5e-112;
  Matches  323;  Conservative   74;  Mismatches  112;  Indels   25;  Gaps    6;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :|||||:: |||:||||:||   |:   |    ||: | |  |  | :: | :|:||
Db        292 EGTLNDCRVMSVDEIFKIEKPG-AHPISFADGKFLRKNDPECDLCGGDPDKTCHMCSCHK 350

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              || ::||: ||:||||:||:||||| |||  || |: |||| |: |:|||| |||||: |
Db        351 CGEKRDPNMQLLCDECNMAYHIYCLSPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKLS 410

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:|||||||||||||||||||||||||: ||||||||:|
Db        411 KKKAKMPSASTESRRDWGRGMACVGRTKECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 470

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:|||||||||:||:|| |: ||||||||::|:||||    | ||

Db        471 VHRPHVGGIHGRSNDGAYSLVLAGGFEDEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 530

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        531 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 590

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              |||||||   | |||||||||||||| || ||| || :| ::| | :||| ||
Db        591 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-------P 643

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              |||  | : : :::|  |: |         |    |  :   : :     || : || |
Db        644 SEKEGKKTKGQSKKQGSEATKRPA-------SDDECPGDSKVLKASDSTDAVEAFQLTPQ 696

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  |||||  | |||:|||||| : |         || |:|::| |:|||||::|:|| |
Db        697 QQRLIREDCQNQKLWDEVLASLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 749

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
              |||||||| |||:||||||||||:|||::| | :|: |||:|   ||||  ||
Db        750 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYVMVLNETLQTLLELFFPGYSKGR 803


RESULT 10
Q96PU4
ID   Q96PU4      PRELIMINARY;      PRT;   802 AA.
AC   Q96PU4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Np95-like ring finger protein (Nuclear zinc finger protein Np97).
GN   Name=NIRF; Synonyms=UHRF2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22165473; PubMed=12176013; DOI=10.1016/S0006-291X(02)00890-2;
RA   Mori T., Li Y., Hata H., Ono K., Kochi H.;
RT   "NIRF, a novel RING finger protein, is involved in cell-cycle
RT   regulation.";
RL   Biochem. Biophys. Res. Commun. 296:530-536(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   PubMed=14741369; DOI=10.1016/S0014-5793(03)01495-9;
RA   Mori T., Li Y., Hata H., Kochi H.;
RT   "NIRF is a ubiquitin ligase that is capable of ubiquitinating PCNP, a
RT   PEST-containing nuclear protein.";
RL   FEBS Lett. 557:209-214(2004).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Davenport J.W., Fernandes E.R., Neale G.A.M., Goorha R.M.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR   EMBL; AB071698; BAB68317.1; -.
DR   EMBL; AF274049; AAM33799.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
```

```
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00184; RING; 2.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Metal-binding; Zinc; Zinc-finger.
SQ   SEQUENCE   802 AA;  89984 MW;  190E26D5A347A7FA CRC64;

  Query Match             59.3%;  Score 1718.5;  DB 2;  Length 802;
  Best Local Similarity   59.2%;  Pred. No. 7.1e-110;
  Matches  316;  Conservative   77;  Mismatches  116;  Indels   25;  Gaps    6;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :|||:|| |||:|||||||   |:   |    |:: | |  |  |  : |  |:| :
Db        291 EGTLNDCKIISVDEIFKIERPG-AHPLSFADGKFLRRNDPECDLCGGDPEKKCHSCSCRV 349

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              |||: :|: ||:||||::|:|||||:|||  || |: |||| |: |:|||| ||||: |
Db        350 CGGKHEPNMQLLCDECNVAYHIYCLNPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKMS 409

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||:|||||||||||||||||: ||||||||:|
Db        410 KKKAKMPSASTESRRDWGRGMACVGRTRECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 469

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQK 238
              |||||| |||||||||:||||||||: |:|| |: |||||||::|:|||| | ||
Db        470 VHRPHVGGIHGRSNDGAYSLVLAGGFADEVDRGDEFTYTGSGGKNLAGNKRIGAPSADQT 529

Qy        239 LTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYK 298
              ||| |||||||| ||::|: |||:::||:||||||:|: || | ||||| ||||||||||
Db        530 LTNMNRALALNCDAPLDDKIGAESRNWRAGKPVRVIRSFKGRKISKYAPEEGNRYDGIYK 589

Qy        299 VVKYWPEKGKS-GFLVWRYLLRRDDDEPGPWTKEGKDRIKKLGLTMQYPEGYLEALANRE 357
              |||||||   | |||||||||||||| || ||| || :| ::| | :||| ||       :
Db        590 VVKYWPEISSSHGFLVWRYLLRRDDVEPAPWTSEGIERSRRLCLRLQYPAGY-----PSD 644

Qy        358 REKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRRTSKKTKVEPYSLTAQ 417
              :| :  | : ::|  |    |          |    || :  : :     :| : || |
Db        645 KEGKKPKGQSKKQPSGTTKRP---------ISDDDCPSASKVYKASDSAEAIEAFQLTPQ 695

Qy        418 QSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQCICCQELVFRPITTVC 477
              |  |||||  | |||:|||: | : |        || |:|::| |:||||||::|:|| |
Db        696 QQHLIREDCQNQKLWDEVLSHLVEGPN-------FLKKLEQSFMCVCCQELVYQPVTTEC 748

Qy        478 QHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQLFPGYGNGR 531
               |||||||| |||:|||||||||:|||::| |  |: |||:|:  ||||  ||
Db        749 FHNVCKDCLQRSFKAQVFSCPACRHDLGQNYIMIPNEILQTLLDLFFPGYSKGR 802


RESULT 11
Q6IP39
ID   Q6IP39      PRELIMINARY;      PRT;   597 AA.
AC   Q6IP39;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LOC432234 protein (Fragment).
GN   Name=LOC432234;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072079; AAH72079.1; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00466; SRA; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
FT   NON_TER     597    597
SQ   SEQUENCE   597 AA;  67054 MW;  273BEC791D9FA86E CRC64;

  Query Match             53.3%;  Score 1544;  DB 2;  Length 597;
  Best Local Similarity   80.4%;  Pred. No. 5.3e-98;
  Matches  270;  Conservative   36;  Mismatches   30;  Indels    0;  Gaps    0;

Qy          2 DSLNDCRIIFVDEVFKIERPGEGSPMVDNPMRRKSGPSCKHCKDDVNRLCRVCACHLCGG 61
              |||||||| ||:|::||| ||      ::| :|::|| ||||||:  | ||:||| :|||
Db        262 DSLNDCRIRFVNEIYKIEEPGSTYLNTESPQKRQNGPECKHCKDNPKRACRMCACCICGG 321

Qy         62 RQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRESKKN 121
              :|||:|||:|||||:|||||| |||| :| :::|||||:|||||||||||||:|:||||
Db        322 KQDPEKQLLCDECDLAFHIYCLKPPLSVIPQDEDWYCPDCRNDASEVVLAGEKLKESKKK 381

Qy        122 AKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESGVHR 181
              |:|||| ||||||||||||::|||||||||||||||:||||:|:||||||||||||||
Db        382 ARMASANSSSQRDWGKGMACVGRSRECTIVPSNHYGPIPGVPVGTLWKFRVQVSESGVHR 441

Qy        182 PHVAGIHGRSNDGSYSLVLAGGYEDDVDHGNFFTYTGSGGRDLSGNKRTAEQSCDQKLTN 241
              |||||||||||||||||||||||||||:|| ||||||||||||||||||||||||:|
Db        442 PHVAGIHGRSNDGSYSLVLAGGYEDDVDNGNEFTYTGSGGRDLSGNKRTAEQSCDQKLSN 501

Qy        242 TNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRNVKGGKNSKYAPAEGNRYDGIYKVVK 301
              |||||||| ||||||:||: |||||:||||||||| || |:||||| ||||||||||||
```

```
Db        502 MNRALALNCSAPINDKEGSIAKDWRAGKPVRVVRNSKGRKHSKYAPEEGNRYDGIYKVVK 561

Qy        302 YWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRIK 337
              |||||||||||||||||||| || ||:|||| : |
Db        562 YWPEKGKSGFLVWRYLLRRDDYEPAPWSKEGKKKKK 597


RESULT 12
Q8C6F1
ID   Q8C6F1      PRELIMINARY;      PRT;   299 AA.
AC   Q8C6F1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male small intestine cDNA, RIKEN full-length
DE   enriched library, clone:2010107E01 product:nuclear protein 95, full
DE   insert sequence. (Fragment).
GN   Name=Uhrf1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Small intestine;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA   Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK075819; BAC35985.1; -.
DR   MGD; MGI:1338889; Uhrf1.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0008283; P:cell proliferation; TAS.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   299 AA;  33647 MW;  512C0893DBABF048 CRC64;

  Query Match           41.7%;  Score 1207;  DB 2;  Length 299;
  Best Local Similarity   73.0%;  Pred. No. 3.8e-75;
  Matches  230;  Conservative   35;  Mismatches   34;  Indels   16;  Gaps    6;

Qy        217 TGSGGRDLSGNKRTAEQSCDQKLTNTNRALALNCFAPINDQEGAEAKDWRSGKPVRVVRN 276
              |||||||||||||| || |||||| ||||||| :||| ::||||:||| |||||||||
Db          1 TGSGGRDLSGNKRTAGQSSDQKLTNNNRALALNCHSPIN-EKGAEAEDWRQGKPVRVVRN 59

Qy        277 VKGGKNSKYAPAEGNRYDGIYKVVKYWPEKGKSGFLVWRYLLRRDDDEPGPWTKEGKDRI 336
              :||||:|||||||||||||||||||:|||||||||||||| || |||:|||||
Db         60 MKGGKHSKYAPAEGNRYDGIYKVVKYWPERGKSGFLVWRYLLRRDDTEPEPWTREGKDRT 119

Qy        337 KKLGLTMQYPEGYLEALANRREKEKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSR 396
              ::|||||||||||||||    || |::   :  | | :| :||| | |      ||:
Db        120 RQLGLTMQYPEGYLEALAN----KEKSRKRPAKALEQGPSSSKTGKSKQK----STGPT- 170

Qy        397 AGSPRRTSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKV 456
              || | |||:|:|||:|: ||::||:||| |||||::|| ||:|    |:|:|||||
Db        171 LSSP-RASKKSKLEPYTLSEQQANLIKEDKGNAKLWDDVLTSLQD-----GPYQIFLSKV 224

Qy        457 EETFQCICCQELVFRPITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPL 516
              :| |||||||||||||:|||||||||||||||||||||||||::|  |   :|||||
Db        225 KEAFQCICCQELVFRPVTTVCQHNVCKDCLDRSFRAQVFSCPACRFELDHSSPTRVNQPL 284

Qy        517 QTVLNQLFPGYGNGR 531
              ||:|||||||:||
Db        285 QTILNQLFPGYGSGR 299


RESULT 13
Q9H6S6
ID   Q9H6S6      PRELIMINARY;      PRT;   189 AA.
AC   Q9H6S6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein FLJ21925.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA   Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA   Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK025578; BAB15177.1; -.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
```

```
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   InterPro; IPR001841; Znf_ring.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
SQ   SEQUENCE   189 AA;  21351 MW;  88C899C7029185AE CRC64;

  Query Match             34.7%;  Score 1005;  DB 2;  Length 189;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-61;
  Matches  189;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        343 MQYPEGYLEALANRREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRR 402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQYPEGYLEALANRREREKENSKREEEEQQEGGFASPRTGKGKWKRKSAGGGPSRAGSPRR 60

Qy        403 TSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQC 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TSKKTKVEPYSLTAQQSSLIREDKSNAKLWNEVLASLKDRPASGSPFQLFLSKVEETFQC 120

Qy        463 ICCQELVFRPITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQ 522
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ICCQELVFRPITTVCQHNVCKDCLDRSFRAQVFSCPACRYDLGRSYAMQVNQPLQTVLNQ 180

Qy        523 LFPGYGNGR 531
              |||||||||
Db        181 LFPGYGNGR 189


RESULT 14
Q8VIA1
ID   Q8VIA1      PRELIMINARY;      PRT;   474 AA.
AC   Q8VIA1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Np95 (Fragment).
GN   Name=np95;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22206529; PubMed=12084726; DOI=10.1074/jbc.M205189200;
RA   Muto M., Kanari Y., Kubo E., Takabe T., Kurihara T., Fujimori A.,
RA   Tatsumi K.;
RT   "Targeted disruption of Np95 gene renders murine embryonic stem cells
RT   hypersensitive to DNA damaging agents and DNA replication blocks.";
RL   J. Biol. Chem. 277:34549-34555(2002).
DR   EMBL; AB066245; BAB79496.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0016567; P:protein ubiquitination; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
FT   NON_TER     474    474
SQ   SEQUENCE   474 AA;  53685 MW;  0EAE841CCD8C7FEA CRC64;

  Query Match             30.6%;  Score 886;  DB 2;  Length 474;
  Best Local Similarity   73.5%;  Pred. No. 8.5e-53;
  Matches  158;  Conservative   22;  Mismatches   27;  Indels    8;  Gaps    1;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPMVDNPMR--------RKSGPSCKHCKDDVNRLCR 52
              |  ||:|||:|||||  || | |  |:: :| :         ||||||: |||| |: ||
Db        260 DSQLNNCRIMFVDEVLMIELPKERRPLIASPSQPPPALRNTGKSGPSCRFCKDDENKPCR 319

Qy         53 VCACHLCGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAG 112
              ||||:|||| |:|||:|||||||||:||| |||:||| | ||||| || |:|||| ||
Db        320 KCACHVCGGREAPEKQLLCDECDMAFHLYCLKPPLTSVPPEPEWYCPSCRTDSSEVVQAG 379

Qy        113 ERLRESKKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRV 172
              |:|:|||| |||||||||:|||||||||||||| ||||||:||:|||||:||||||||||
Db        380 EKLKESKKKAKMASATSSSRRDWGKGMACVGRTTECTIVPANHFGPIPGVPVGTMWRFRV 439

Qy        173 QVSESGVHRPHVAGIHGRSNDGSYSLVLAGGYEDD 207
              |||||||| |||||  |||||:||||||||||||
Db        440 QVSESGVHRLHVAGIPWRSNDGAYSLVLAGGYEDD 474


RESULT 15
Q8BJP6
ID   Q8BJP6      PRELIMINARY;      PRT;   516 AA.
AC   Q8BJP6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length
DE   enriched library, clone:B430210O11 product:hypothetical Ubiquitin
DE   domain containing protein, full insert sequence.
GN   Name=Uhrf2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
```

```
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Adipose;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK080925; BAC38081.1; -.
DR   HSSP; Q9UIG0; 1F62.
DR   MGD; MGI:1923718; Uhrf2.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011011; FYVE_PHD_ZnF.
DR   InterPro; IPR003105; SRA_YDG.
DR   InterPro; IPR000626; Ubiquitin.
DR   InterPro; IPR001965; Znf_PHD.
DR   Pfam; PF00628; PHD; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   Pfam; PF02182; YDG_SRA; 1.
DR   SMART; SM00249; PHD; 1.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50016; ZF_PHD_2; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   516 AA;  57761 MW;  3D6862E9D1E1DEDD CRC64;

  Query Match             28.2%;  Score 816.5;  DB 2;  Length 516;
  Best Local Similarity   67.1%;  Pred. No. 5.7e-48;
  Matches  141;  Conservative   31;  Mismatches   35;  Indels    3;  Gaps    2;

Qy          1 DDSLNDCRIIFVDEVFKIERPGEGSPM--VDNPMRRKSGPSCKHCKDDVNRLCRVCACHL 58
              : :|||||:: |||:||||:||   |:   |    ||: | |  |  | :: | :|:||
Db        292 EGTLNDCRVMSVDEIFKIEKPG-AHPISFADGKFLRKNDPECDLCGGDPDKTCHMCSCHK 350

Qy         59 CGGRQDPDKQLMCDECDMAFHIYCLDPPLSSVPSEDEWYCPECRNDASEVVLAGERLRES 118
              || ::||: ||:||||:||:||||| |||  || |: |||| |: |:|||| |||||: |
Db        351 CGEKRDPNMQLLCDECNMAYHIYCLSPPLDKVPEEEYWYCPSCKTDSSEVVKAGERLKLS 410

Qy        119 KKNAKMASATSSSQRDWGKGMACVGRTKECTIVPSNHYGPIPGIPVGTMWRFRVQVSESG 178
              || ||| ||:: |:||||:||||||||||||||||||||||||: ||||||||||:|
Db        411 KKKAKMPSASTESRRDWGRGMACVGRTKECTIVPSNHYGPIPGIPVGSTWRFRVQVSEAG 470

Qy        179 VHRPHVAGIHGRSNDGSYSLVLAGGYEDDV 208
              |||||| |||||||||:||||||||:||:|
Db        471 VHRPHVGGIHGRSNDGAYSLVLAGGFEDEV 500


Search completed: July 25, 2005, 17:06:45
Job time : 183 secs
```